Water Vapour Climate Change Initiative (WV_cci) - Phase One

ATBD Part 2 - IMS L2 Product

Ref: D2.2

Date: 27 March 2019

Issue: 1.0

For: ESA / ECSAT

Ref: CCIWV.REP.005

This Page is Intentionally Blank

Project : **Water Vapour Climate Change Initiative (WV_cci) - Phase One**

Document Title: **ATBD Part 2 - IMS L2 Product**

Reference : **D2.2**

Issued : **27 March 2019**

Issue : **1.0**

Client: **ESA / ECSAT**

Author(s) : Richard Siddans

Document Change Log

Issue/ Revision	Date	Comment
1.0	27 March 2019	Initial issue

TABLE OF CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

This Page is Intentionally Blank

1. INTRODUCTION

1.1 Purpose

This document is the algorithm theoretical baseline document (ATBD) for the RAL Infrared and Microwave Sounder (IMS) scheme, as used within the ESA CCI Water Vapour project.

The purpose of the ATBD is to provide detailed mathematical and physical descriptions of this algorithm, along with an overview of algorithm inputs and outputs, data products, algorithm validation and error analysis

1.2 Scope

This document describes the IMS algorithm which has been applied to generate the data used in the first phase of the project. The algorithm was developed prior to the CCI project via a Eumetsat study and funding from the UK National Centre for Earth Observation (NCEO).

While this document provides an overview of product quality from previous work, note that these topics will be covered in more depth by other CCI documents including the E3UB and PVIR.

2. ALGORITHM DEFINITION

2.1 Introduction

The Infrared Microwave Sounding (IMS) scheme employs the optimal estimation method (OEM) to jointly retrieve water vapour, temperature and ozone profiles surface spectral emissivity and cloud parameters from the Metop sounding instruments IASI, MHS and AMSU.

2.2 Heritage

The IMS scheme was developed in the context of a Eumetsat study (RD-1) which started with the specification of the Eumetsat operational optimal estimation scheme used for the version 6 operational product (RD-3).

The Eumetsat version 6 product is generated using a series of algorithms, including a piece-wise linear regression (PWLR) scheme which uses measurements by IASI, AMSU and MHS together to predict water vapour, temperature and ozone profiles. This PWLR is based on regressing observed radiances from the three sounders against ECMWF analyses. In cloud free scenes, output from the PWLR is used as *a priori* information for an OEM retrieval from IASI only. PWLR results are directly reported in the L2 product, as well as results from the OEM (for the sub-set of cloud-free scenes). Separate schemes are used to perform cloud flagging and to retrieve trace-gases (e.g. CO). It is specifically the temperature and water vapour OEM algorithm which was used as a basis for the RAL IMS development.

During the study three advances over the Eumetsat OEM were developed:

- Information from IASI, AMSU and MHS measurements are combined
- Spectral emissivity is jointly retrieved
- Cloud parameters are included in the retrieval, enabling the scheme to be applied to cloudy scenes.

Those three extensions were found to improve agreement with ECMWF analyses of lower tropospheric water vapour and to reduce the sensitivity to small amounts of cloud contamination; significantly improving the coverage of useful data.

The IMS scheme was subsequently developed through work within the UK National Centre for Earth Observation (NCEO). The scheme now uses a weak prior constraint, based on zonal mean climatology. It is in practice therefore independent of ECMWF analyses or re-analyses. The IMS scheme uses the RTTOV (v10) radiative transfer

model to simulate brightness temperature observations of the IASI and microwave sounders.

Via NCEO, the IMS scheme has been applied to process the complete IASI Metop A mission from 2007 to 2016. This “Version-1” dataset (RD-9) is made available in Year-1 to the CCI+ Water Vapour project. The data is now archived at CEDA (http://www.ceda.ac.uk/).

Since the generation of this dataset, the IMS scheme has developed further, building on new capabilities introduced in RTTOV 12 as follows:

- The quality of the height-resolved ozone retrieval has been much improved.
- Joint retrieval of additional trace gases including carbon monoxide (CO), nitric acide (HNO3), methanol (CH3OH), ammonia (NH3) and formic acide (HCOOH).
- Joint retrieval of dust and sulfuric acid aerosol optical depth.
- Cloud is now represented as a multiple scattering layer using spectral optical properties of ice / liquid cloud. Effective radius, altitude and optical depth of the the cloud are now jointly retrieved.

Full mission processing with the improved scheme has yet to be carried out and this ATBD focuses on the version 1 scheme used to generate the data to be used in CCI water vapour phase 1.

2.3 Instruments

2.3.1 Metop

The IMS scheme uses the three sounding instruments, IASI, AMSU and MHS, on-board the Metop platform of the Eumetsat Polar System (EPS). There are three Metop plaforms: Metop-A, -B and –C were launched respectively on 19 October 2006, 17 September 2012 and 7 November 2018. Data from Metop-A and B IASI is available from 29 May 2007 and 20 February 2013, respectively. All relevant sensors on both platforms are still functioning.

2.3.2 Infrared Atmospheric Sounding Interferometer (IASI)

Metop IASI (RD-4) is a nadir viewing infra-red Fourrier transform spectrometer which provides spectra at 0.5 cm^{-1} apodised resolution, sampled every 0.25 cm^{-1}, from 625 to 2760 cm^{-1}. Spectra are measured with 4 detectors, each with a circular field of view on the ground (at nadir) of approximately 12 km diameter, arranged in a 2 x 2 grid within a 50 x 50 km^2 field-of-regard (FOR). IASI scans to provide 30 fields-of-regard (FOR) (i.e. 120 individual spectra) evenly distributed across a 2200 km wide swath.

2.3.3 AMSU and MHS

The ATOVS (Advanced TIROS (Television and Infrared Observational Satellite) Operational Vertical Sounder) is a sounding instrument package flown on operational polar orbiting platforms since NOAA-15 (launched May 1998). On Metop, it comprises the Advanced Microwave Sounding Unit A, the Microwave Humidity Sounder (MHS) and the High Resolution InfraRed Sounder (HIRS/4; not used in the IMS scheme).

AMSU-A and MHS are across-track scanning microwave radiometers. AMSU-A measures in 15 spectral bands, including coverage of the 50GHz oxygen line, used for temperature sounding. MHS adds five channels around the 183 GHz water vapour line. The spectral coverage is summarised in Table 2-1 and illustrated in Figure 2-1.

Both the microwave sounders have a scan range of around ±48° from nadier, giving a swath width similar to that of IASI. While the instantaneous fields of view (IFOV) of the sounders differ, the scans are synchronised to give a systematic regular sampling pattern from the four instruments, as illustrated in Figure 2-2. The IFOV of AMSU (at nadir) is circular with a diameter of around 48km. 30 of these footprints are measured in each scan, closely matching the 30 FOR of IASI (i.e. there are 4 IASI measurements within each AMSU IFOV). MHS has a circular footprint of around 16km diameter at nadir. Three scans of MHS are made for each scan of IASI/AMSU (yielding 9 MHS measurements for each AMSU IFOV).

Channel	Channel frequency / GHz	Nominal bandwidth / MHz	Noise Equivalent Brightness Temperature / K
1	23.8	270	0.3
2	31.4	180	0.3
3	50.3	180	0.4
4	52.8	400	0.25
5	53.59±0.115	170	0.25
6	54.4	400	0.25
7	54.94	400	0.25
8	55.5	330	0.25
9	F_{LO}= 57.290344	330	0.25
10	F_{LO} ± 0 .217	78	0.4
11	F_{LO} ± 0 .3222 ± 0.048	36	0.4
12	F_{LO} ± 0 .3222 ± 0.22	16	0.6
13	F_{LO} ± 0 .3222 ± 0.01	8	0.8
14	F_{LO} ± 0 .3222 ± 0.0045	3	1.2
15	89	<6000	0.5
16 (MHS H1)	89	2800	1
17 (MHS H2)	157	2800	1
18 (MHS H3)	183.311±1.00	500	1
19 (MHS H4)	183.311±1.00	1000	1
20 (MHS H5)	190.311	2200	1

Table 2-1: AMSU (1-15) and MHS (16-20) channel characteristics. For the AMSU channels 9-14, F_{LO} indicates the local oscillator frequency about which the spectral response is indicated (from RD-5).

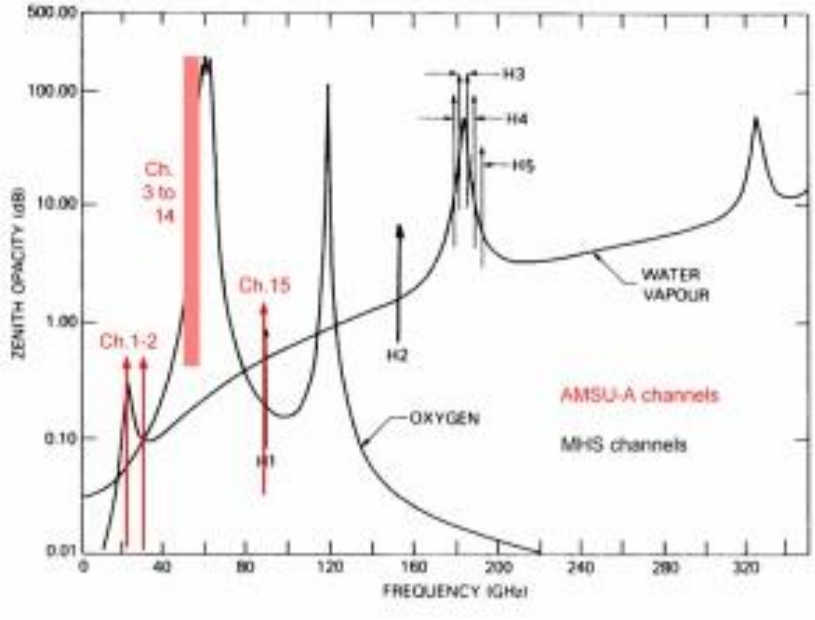

Figure 2-1: Illustration of the spectral coverage of AMSU and MHS channels (from Eumetsat web-site).

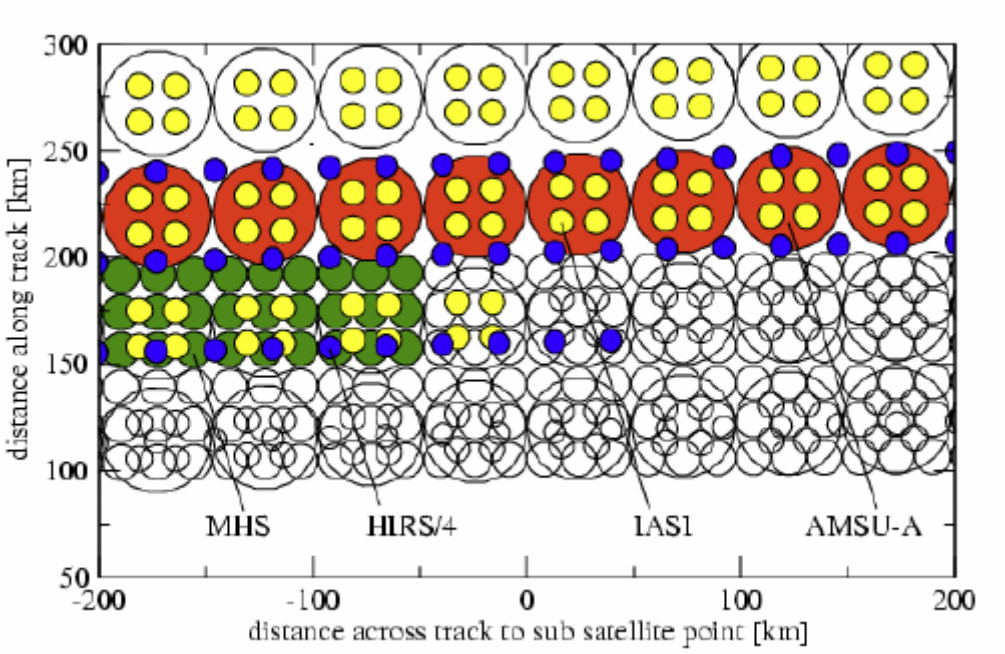

Figure 2-2: Illustration of the footprints of Metop AMSU (red), MHS (green), IASI (yellow) and HIRS/4 (blue) from RD-6.

2.4 The Optimal Estimation Method

The scheme is based on the optimal estimation method (OEM, RD-7), which solves an otherwise under-constrained inverse problem by introducing prior information. The method finds the optimal state vector $\boldsymbol{x}$ (which contains the parameters we wish to retrieve) by minimising the following cost function:

$$\chi^2 = \left(\boldsymbol{y} - F\left(\boldsymbol{x}\right)\right)^T \boldsymbol{S_y^{-1}}\left(\boldsymbol{y} - F\left(\boldsymbol{x}\right)\right) + \left(\boldsymbol{a} - \boldsymbol{x}\right)^T \boldsymbol{S_a^{-1}}\left(\boldsymbol{a} - \boldsymbol{x}\right)$$

Equation 1

where $\boldsymbol{y}$ is a vector containing each measurement used by the retrieval; $\boldsymbol{S_y}$ is a covariance matrix describing the errors on the measurements; $\boldsymbol{F}(\boldsymbol{x})$ is the *forward model* (FM), which predicts measurements given $\boldsymbol{x}$; $\boldsymbol{S_a}$ is the *a priori* covariance matrix, which describes the assumed errors in the *a priori* estimate of the state, $\boldsymbol{a}$. Minimising this cost function, corresponds to maximising the probability of the solution state being correct, given the uncertainties on the measurements and prior state as expressed by their respective covariances.

If the FM is linear then the solution is given by:

$$\hat{\boldsymbol{x}} = \boldsymbol{a} + \left(\boldsymbol{K^T S_y^{-1} K} + \boldsymbol{S_a^{-1}}\right)^{-1} \boldsymbol{K^T S_y^{-1}} \left(\boldsymbol{y} - F(\boldsymbol{a})\right)$$

Equation 2

where $\boldsymbol{K}$ is the weighting function matrix, the rows of which contain the derivatives of the FM with respect to each element, *j*, of the state vector:

$$\boldsymbol{K_j} = \frac{\partial\, F(\boldsymbol{x})}{\partial\, \boldsymbol{x_j}}$$

Equation 3

In practice, the FM is non-linear and the solution state needs to be found iteratively, starting from a first guess state, $\boldsymbol{x_0}$. We adopt the well known Levenburg-Marquardt method (RD-8) to find the solution by iterating the following equation:

$$\boldsymbol{x_{i+1}} = \boldsymbol{x_i} + \left(\boldsymbol{K_i^T S_y^{-1} K} + \boldsymbol{S_a^{-1}} + \gamma \boldsymbol{I}\right)^{-1} \left[\boldsymbol{K_i^T S_y^{-1}} \left(\boldsymbol{y} - F(\boldsymbol{x_i})\right) - \boldsymbol{S_a^{-1}}(\boldsymbol{x_i} - \boldsymbol{x_a})\right]$$

Equation 4

Where γ is a parameter which controls the magnitude of the state vector update at each iteration, i. $\boldsymbol{K_i}$ is the weighting function evaluated for the current estimate of the state, $\boldsymbol{x_i}$. If γ is negligible and the FM is linear (within the relevant range), then the iteration will immediately find the cost function minimum. However, if the FM is non-linear the iteration will not necessarily lead to the minimum and may even result in a higher cost function value. If γ is large, the iteration will tend to follow the gradient of the cost function, and this should always lead to a reduction in cost. Convergence towards the minimum will however be relatively slow. By controlling the magnitude of γ during the iteration, efficient convergence can usually be assured. Here we adopt the following approach:

1. Initially, γ is (somewhat arbitrarily) set equal to 0.001

2. First guess, $\boldsymbol{x_i}$, (for i=0) is defined. The cost function for this state, χ_i^2, is evaluated.

3. $\boldsymbol{x_{i+1}}$ is estimated, using Equation 4 and the new cost function value is determined, χ_{i+1}^2. The iteration proceeds according to the following logic:

    - If the absolute difference between χ_{i+1}^2 and χ_i^2 is smaller than a given "convergence threshold" (usually 1) is then the scheme proceeds to step 4.
    - If χ_{i+1}^2 is larger than χ_i^2 (the updated state is worse than the previous estimate), then γ is increased by a factor of 10 and the retrieval returns to the beginning of step 3, starting again $\boldsymbol{x_i}$. I.e. iteration count i is not incremented.
    - If χ_{i+1}^2 is smaller than χ_i^2 (the updated state is better than the previous estimate), then γ is decreased by a factor of 10 and the retrieval returns to the beginning of step 3, starting from the updated (better) state, i.e at this point iteration count i is incremented.

4. At this point the last iteration will have resulted in a change of cost which is smaller than the convergence threshold. This could indicate the cost function minimum has been found but may also arise if γ has become very large. A further (potentially final) iteration of the retrieval is now carried out applying Equation 4 with γ=0, to give state $\boldsymbol{x_n}$ and cost χ_n^2. If again the cost function change is smaller than the convergence threshold then the retrieval finishes the final state $\boldsymbol{x_n}$ is reported as the solution. If the cost change is larger than the convergence threshold, then γ is reset to its initial value (0.001) and the retrieval "re-starts" from the state ($\boldsymbol{x_i}$ or $\boldsymbol{x_n}$) which gave the lowest cost.

   The number of allowed iterations and attempted "re-starts" from false minima are subject to pre-defined limits. If these are exceeded during the iteration approach above, then the retrieval is stopped. Results are still reported, with a flag indicating that

convergence has not been fully achieved. These results may still be useful (especially in the case that the solution cost function value is low), but must be used with caution.

Note iteration count $\boldsymbol{i}$ is only updated when a state vector update gives rise to an improvement in the cost. The final value of $\boldsymbol{i}$ is reported as the number of iterations, $\boldsymbol{N\_i}$. The total number of retrieval "steps" (i.e. any state vector update, good or bad), $\boldsymbol{N_s}$, is also usually recorded for information (e.g. to enable optimisation of computational cost by identifying scenes requiring a large number of evaluations of the FM).

The retrieval final cost function value at is always stored, so it can be used to quality-control retrievals later. For detailed analysis, the fit *residual spectum* ($\boldsymbol{y} - F(\boldsymbol{x})$ at solution) is also required.

2.5 Measurements

2.5.1 IASI

The version 1 IMS scheme is based on the operational Eumetsat OEM for IASI and implements the same approach to define the IASI measurement vector and associated covariance ($\boldsymbol{S_y}$). This defines (i) some pre-processing of IASI observations to reduce noise and filter instrument artefacts; (ii) a specific sub-set of IASI channels to be used (determined via an information-content analysis); (iii) a scan angle dependent bias correction; (iv) a specific, fixed measurement error covariance matrix ($\boldsymbol{S_y}$) intended to reflects a combination of instrumental and forward model errors:

2.5.1.1 Pre-processing of L1C spectra

IMS uses IASI measured (L1C) spectra which have been compressed and re-constructed using the operational principal components (RD-14). This has the effect of reducing noise on the sub-set of channels which are used in the scheme. A further filter is applied to remove some instrumental artefacts, described in RD-11.

2.5.1.2 Channel selection

139 IASI channels were selected by EUMETSAT. The centre wavenumbers of these channels aare as follows:

662.5, 667, 667.75, 668.25, 668.75, 670.25, 680, 680.75, 693, 697.75, 704.5, 705.5, 717.75, 719.25, 720.25, 721.5, 723, 726.25, 729.25, 731.25, 741.25, 742, 744.5, 746,

752.25, 756.75, 769, 781.25, 791.75, 796.5, 798.5, 798.75, 811.75, 814.5, 831.75, 835.5, 852.75, 879.25, 903.25, 955.25, 957, 979.75, 1006.25, 1010.75, 1028.5, 1034.75, 1039.75, 1043.75, 1045, 1052.25, 1054.75, 1060.25, 1064.5, 1066.25, 1069, 1072.5, 1074.5, 1115.75, 1121, 1135.75, 1136.5, 1141.5, 1148.5, 1157, 1182.25, 1185.75, 1186.25, 1197, 1201.5, 1204.25, 1206.25, 1210.5, 1211.25, 1211.75, 1214, 1216.5, 1220.25, 1221, 1224.25, 1225.5, 1230, 1234.25, 1239.25, 1240.5, 1242.75, 1244.25, 1245.25, 1247.5, 1252, 1257.75, 1260.25, 1266.25, 1268.25, 1271, 1278.5, 1282.5, 1284, 1285.5, 1288.75, 1291, 1293, 1293.5, 1296.75, 1300, 1304.75, 1305.75, 1306.75, 1313.5, 1316.25, 1317.5, 1325.75, 1331, 1335, 1335.75, 1337.75, 1338.25, 1339, 1340.25, 1345, 1360, 1361.25, 1361.75, 1373, 1381.25, 1381.75, 1395.25, 1399.25, 1400.25, 1401, 1407.75, 1408.5, 1456.75, 1465.75, 1472.75, 1539.75, 1566.75, 1726, 1819.5, 1900cm^{-1}.

The channels positions are illustrated in Figure 2-3.

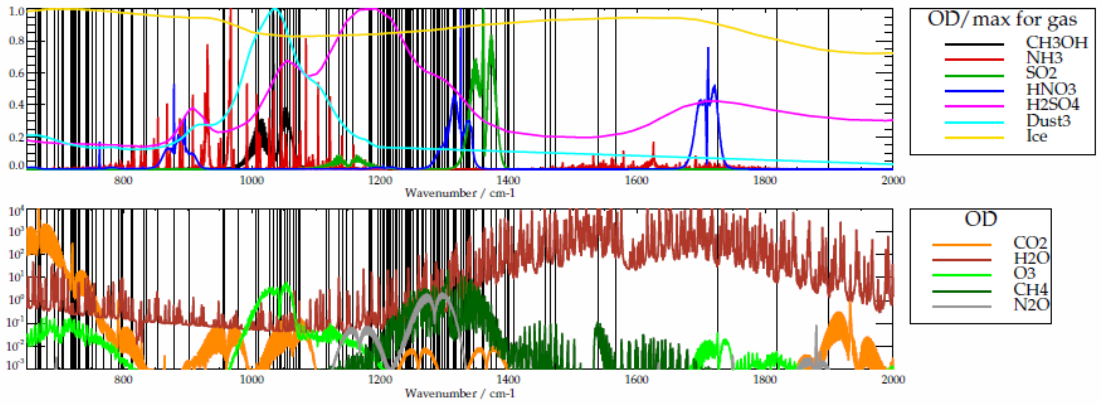

Figure 2-3: IASI channels used by the IMS scheme (indicated by vertical bloack lines). Top panel shows the relative optical depth of some minor trace-gases, together with those of desert dust, sulphuric acid aerosol and ice cloud. These spectra are normalised by their peak optical depth in the full IASI spectral range). The bottom panel shows (not normalised) nadir optical depth of major absorbers. Trace-gas optical depth spectra are taken from RD-15.

2.5.1.3 Bias correction and measurement covariance

In the original Eumetsat scheme, a bias correction for IASI was determined by by comparing simulated spectra based on PWLR profiles to observed radiances for a represtative subset of IASI scenes. This gave a bias correction parameterised in terms of two vectors, $\boldsymbol{b_0}$ and $\boldsymbol{b_1}$, which define a fixed bias spectrum and a scan dependent term. The measurement used in the retrieval, $\boldsymbol{y}$, is then given by

$$\boldsymbol{y} = \boldsymbol{y''} - \boldsymbol{b_0} - \boldsymbol{b_1}\,(\sec(\theta) - \mu_0)$$

Equation 5

Where $\boldsymbol{y''}$ is the double-filtered measurement vector (see above); θ is the satellite zenith angle (at the ground) and μ_0 is a constant.

In the IMS scheme the same vectors $\boldsymbol{b_0}$ and $\boldsymbol{b_1}$ are used, however the weights used in any given scene are retrieved parameters (see section 2.7 below) and the bias correction is instead applied to the simulated measurement:

$$F(\boldsymbol{x}) = R(\boldsymbol{x}) - x_{b0}\boldsymbol{b_0} - x_{b1}\boldsymbol{b_1}$$

Equation 6

Where x_{b0} and x_{b1} are the state vectors element which the two bias correction patterns and $R(\boldsymbol{x})$ denotes the result for RTTOV, which depends on all other elements of the state vector.

2.5.1.4 Measurement covariance

The estimated measurement covariance was also determined by Eumetsat from statistics of FM simulations, based on PWLR from the bias corrected measurements. The bias correction patters and the measurement covariance are illustrated in Figure 2-4.

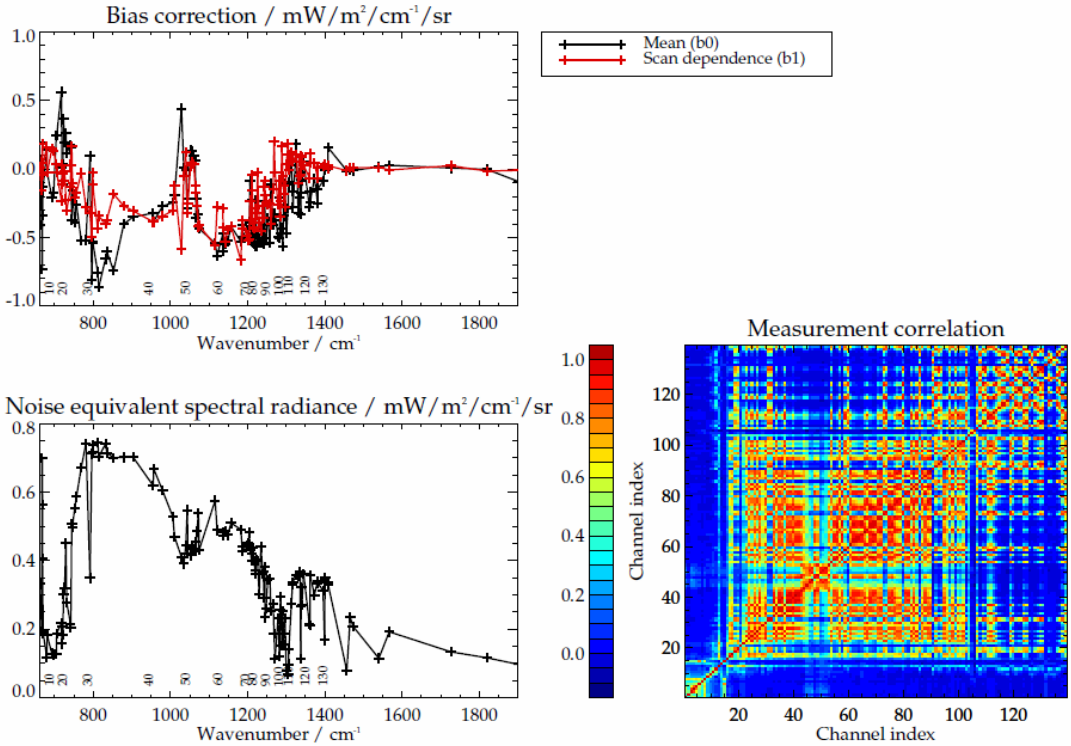

Figure 2-4: Illustration of the IASI bias correction and measurement error covariance. Top left-hand panel shows the bias correction parameters. Numbers above the x-axis give the channel index from the 139 channel sub-set.. Bottom left shows the estimated noise-equivalent spectral radiance, i.e. the square-root diagonal elements of the measurement covariance matrix. Bottom right shows the measurement correlation matrix.

2.5.2 AMSU and MHS

Normally IMS uses all 20 (AMSU-A+MHS) microwave channels. However, channels 7 and 8 on Metop A both exhibited problems during the mission and these have been completely excluded from the version 1 processing. The impact of excluding these channels was checkd in the Eumetsat study and found to be very small.

An across track-bias dependent bias correction is applied to the AMSU and MHS measurements. This was derived during the Eumetsat study (RD-1) as follows:

- The RAL implementation of the original IASI only OEM scheme was applied to process three days of Metop B data (17 April, 17 July, 17 October 2013).
- The resulting profiles were then used to simulate radiances in the AMSU and MHS channels over strictly cloud-free sea, between 60S and 60N (to avoid ice covered sea). The sea surface emissivity was modelled using the model internal to RTTOV. The IASI (v6) cloud mask was used to define cloud-free scenes.
- Differences between simulations and observations were are averaged into six regularly spaced cross-track bins (indexed by the 30 IASI/AMSU IFOVs).

- The resulting mean differences are interpolated in IFOV index to obtain the bias correction applicable to a given scene.

The observation covariance assumed in the retrievals ($\boldsymbol{S_y}$) was also derived by computing the covariance of differences between simulations and the bias corrected measurements, using the same set of cloud-free scenes over sea.. A 3-sigma test was applied to exclude outliers.

The bias correction and measurement covariance is illustrated in Figure 2-5.

Note that:

- Computed in this way, the bias correction and the measurement error covariance reflect a combination of errors in the measurements themselves and in the assumed forward model.
- Although he bias correction and measurement errors are computed from scenes over cloud-free sea only, but they are used for observations over both land and sea and in cloudy conditions. I.e. it is implicitly assumed that instrument and forward model errors are similar over both land and sea and in cloudy scenes.
- IASI measurement errors are assumed to be uncorrelated with those of AMSU and MHS.

2.5.3 Selection of scenes

In the version 1 processing, a simple brightness temperature difference test is applied to detect optically thick and/or high altitude cloud. This selection is based on the difference in brightness temperature between the IASI observation at 950 cm-1 and that simulated on the basis of the first guess state, which (as described below) is based on ECMWF analysis. If this difference (observation – simulation) is outside the range of -5 to 15K, the the scene is not processed.

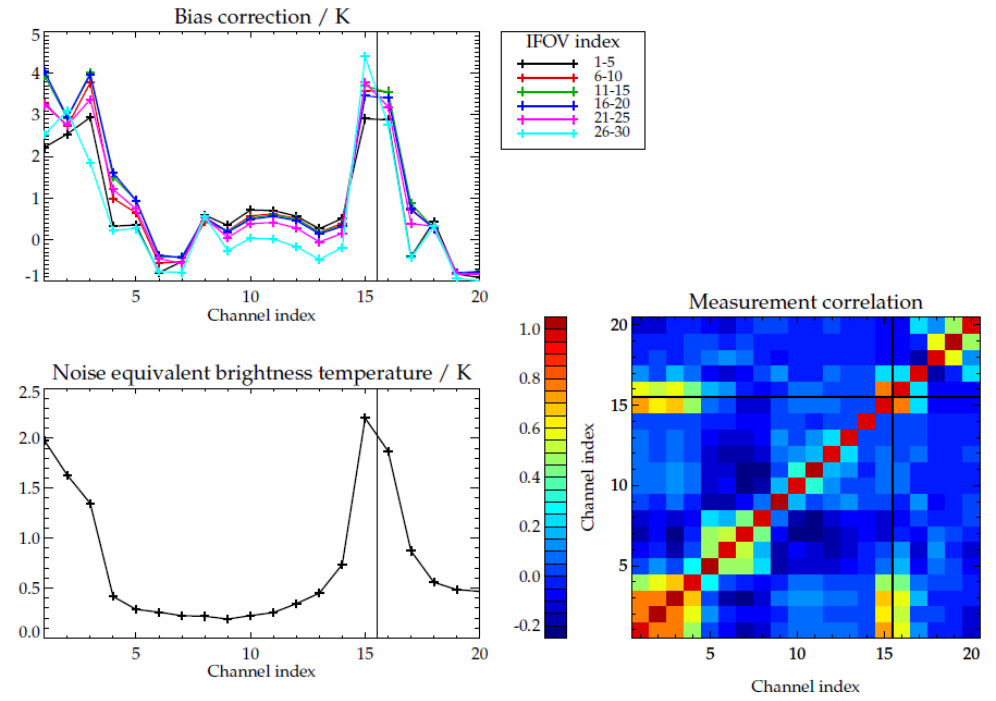

Figure 2-5: Illustration of the AMSU (channel index 1-15) and MHS (16-20) bias correction and measurement error covariance. Top left-hand panel shows the derived bias correction (top left) as a function of across-track IFOV index. Bottom left shows the estimated noise-equivalent brightness temperature (in K), i.e. the square-root diagonal elements of the measurement covariance matrix. Bottom right shows the measurement correlation matrix.

2.6 Forward model

RTTOV version 10.2 (Matricardi 2009) performs is the basis of the FM. RTTOV estimates radiances convolved with the IASI spectral response function by use of spectrally-averaged layer transmittances, based on a fixed set of coefficients which weight atmospheric-state-dependent predictors. The model is sufficiently fast to enable global processing of the IASI mission with modest computational resources.

The version 1 scheme uses the following RTTOV coefficient files (obtained from the NWP SAF website):

- AMSU and MHS: Version 7 coefficients on 54 levels created on 16 May 2013.
- IASI: Version 9 coefficients on 101 levels created on 28 March 2013

With version 7 coefficients, RTTOV can simulate variations in water vapour and ozone (sufficient for modelling the MW channels). Most other trace-gases with significant absorption are included in the modelled transmissions assuming a fixed profile. Version 9 coefficients enables RTTOV to also simulate variations in carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O) and carbon monoxide (CO). The IMS scheme

models the variations of the greenhouse hases (CO_2, CH_4 and N_2O) using a monthly latitude dependent climatology derived from MACC flux inversions. Version 1 of IMS does not use spectral channels affected with significant absoprtion by CO.

Profiles of water vapour, and ozone and temperature are defined by the retrieval state vector (see below). These are defined on the 101 pressure levels on which the RTTOV 9 coefficients are given (for IASI). Profiles are interpolated from this grid to the coarse 54 level grid used for the MW coefficients. Surface temperature and surface emissivity are also defined by the state vector. 2m temperature and 2m water vapour (also input parameters to RTTOV) are defined by interpolating the profiles defined by the state vector.

Surface pressure is defined from ECMWF analysis (ERA interim, RD-17), adjusted to the mean altitude within the IASI footprint assuming the logarithm of the surface pressure varies linearly with the difference between the IASI altitude and that of the ECMWF model. This is the only parameter defined directly from NWP data in the IMS version 1 data.

RTTOV is capable of simulating cloud either as scattering and absorbing profiles (defined in terms of profiles of cloud cover, liquid and ice water content) or as a simple opaque black body at defined altitude, occupying a defined area fraction of the scene. The simple black body model is adopted in the version 1 scheme.

2.7 State vector

IMS jointly retrieves a number of products, which are represented in the single state vector $\boldsymbol{x}$. Each of the products are defined individually below. Products are defined such that the a priori covariance is diagonal i.e. no correlations are assumed either between the elements corresponding to an individual product or between the different products.

2.7.1 Temperature, water vapour and ozone profiles

In IMS, water vapour temperature profiles and ozone profiles are internally represented using basis functions, $\mathbf{M_x}$, which are the Eigenvectors of an assumed covariance matrix which represents the prior variability of the profile on the 101 RTTOV pressure levels. 28 vectors are fitted for water vapour and 18 for water vapour and 10 for ozone. In the original Eumetsat scheme the *a priori* state was defined from the PWLR and

covariances were derived from differences between PWLR and ECMWF analysis. In IMS a weaker, “climatological” prior constraint is used: The covariance matrices are computed using ECMWF analyses for the 3 test days of the Eumetsat study (17 April 17 July, 17 October 2013). The zonal mean over all 3 days is computed. The covariance is then calculated using the differences between all the individual ECMWF profiles and the zonal mean state (linearly interpolated to each latitude). I.e. the covariance of departures from the 3 day zonal mean is computed: The assumed mean state varies with latitude but a single covariance applicable globally is derived. For temperature the mean state and covariances are computed in K. For water vapour and ozone they are computed using the logarithm of the volume mixing ratio in ppmv. Using the Eigenvectos of this covariance, temperature profiles in (K) on the 101 RTTOV pressure levels are defined from the corresponding 28 elements of the state vector as follows:

$$\boldsymbol{T} = \boldsymbol{m_T}(\lambda) + \boldsymbol{M_T x_T}$$

Equation 7

Where $\boldsymbol{m_T}$ is the mean profile, interpolated to the latitude of a specific observation; $\boldsymbol{M_T}$ is the marix of Eigenvectors of the covariance matrix and $\boldsymbol{x_T}$ contain the relevant sub-set of state vector.

Water vapour and ozone profiles (in ppmv) are defined similarly (now with exponent):

$$\boldsymbol{w} = e^{\boldsymbol{m_W}(\lambda)+\boldsymbol{M_W x_W}}$$

Equation 8

In terms of the state vector representation used in the OEM algebra, the *a prior* state vector is all zero (the mean profile is added in the FM). The prior covariance is diagonal with variances given by the Eigenvalues of the covariance matrix.

In order to speed up convergence the first guess state is estimated using the ERA-interim analysis profiles. This implemented by performing a linear optimal estimation fit of the state vector elements to the ERA-interim profile, i.e using the ERA-interim profile as “measurement” and Equation 7 as forward model. The prior state is as defined above (a zero vector and diagonal covariance). This ensures the first guess state is consistent with the way profiles are represented in the full retrieval. In this way the impact of using NWP analysis to defined the first guess is only to reduce the number of iterations required to achieve convergence; the approach does not significantly affect the retrieval solution.

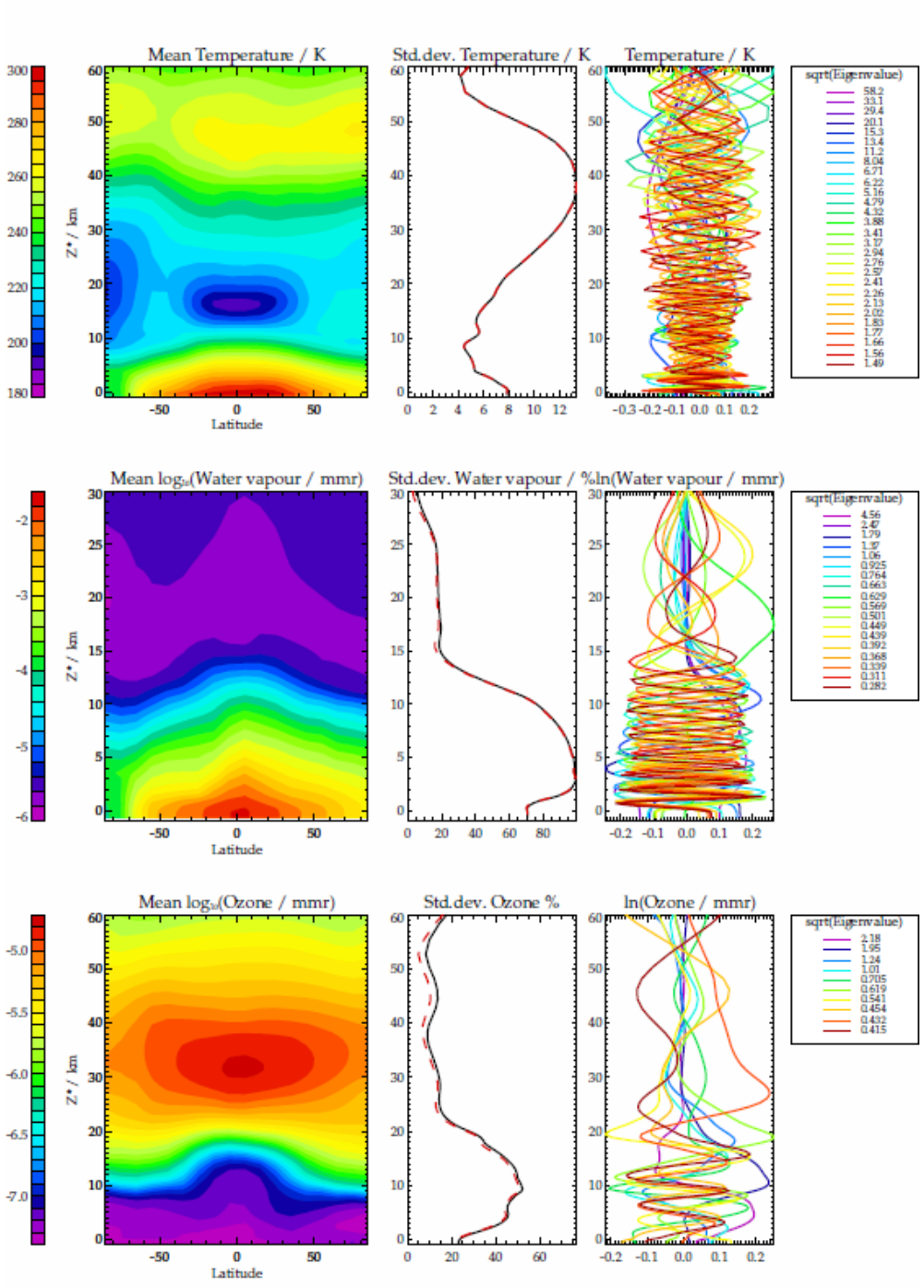

Figure 2-6: Illustration of the climatology used to define the prior state and covariance for temperature (top), water vapour (middle) and ozone (bottom). Left hand panels show the zonal mean profiles. Centre panels show the standard deviation of individual profile departures from the zonal mean. Right-hand panels show the Eigenvectors of the covariance matrix. The legend gives the square-root of the corresponding Eigenvalue.

2.7.2 Surface Temperature

The surface temperature prior state is defined similarly to the temperature profile: The zonal mean surface temperature derived from ECMWF analysis on the 3 days is used as the prior value and the error is given by the standard deviation of departures from the zonal mean (accumulated globally). The first guess value is taken from the ERA-interim analysis.

2.7.3 Surface Emissivity

Emissivity has been included in the state vector in terms of eigenvectors, analogous to the approach use to represent the atmospheric profiles. The basis of the approach is to construct the covariance of the global variation of spectral emissivity in the MW and IR channels using the RTTOV emissivity atlases. RTTOV has separate built-in atlases of emissivity for the MW and IR channels. These are combined to obtain a joint spectral spectral convariance by running RTTOV to generate emissivity in all the channels used by IMS for all IASI scenes (land and sea) for the three test days used in the study (17 April, 17 July, 17 October 2013). The eigenvectors and values of this covariance are obtained and used as the basis of the set of spectral patterns used, and the associated *a priori* covariance. Because RTTOV provides co-located emissivity spectra for both MW and IR ranges, the principle components include correlations between MW and IR, which enables IASI measurement to constrain the emissivity used in the MW and vice-versa. Also because the values span global variation in a realistic manner, the eigenvalues provide suitable values to use as diagonal elements of the *a priori* covariance for emissivity (off diagonal elements being zero). However, this approach is not sufficient, because only a limited amount of spectral information is represented in the RTTOV atlases. To accurately simulate spectra in all used IASI channels, it is necessary to introduce further spectral patterns (see below).

Note that in RTTOV MW land emissivity can come from two atlases, either TELSEM (RD-18) or CNRM (RD-19). Sea emissivity is calculated using the FASTEM model independent of the selection of land atlas. TELSEM is based on SSMI observations (and is the default setting for RTTOV v10.2). The CNRM atlas is based on AMSU A and B. In the version 1 scheme emissivity patterns are based on the TELSEM Atlas.

Both MW databases are interpolated in RTTOV to define vales in the strong absorbing channels from (measurements of emissivity are only available in the window channels). The spectral interpolation is (probably) of little importance for the simulation of radiances, but is found to introduce some (presumably numerical) artefacts into the spectral emissivity eigenvectors. We therefore simplify the representation of spectral emissivity in the MW before computing the covariance and eigenvectors as follows:

- AMSU channels 1-3 and 15 are taken as provided by RTTOV.

- AMSU channels 4-14 (all in range 52.8-57.3 GHz) are assumed to have the same emissivity as channel 3 (50.3 GHz)
- MHS channel 1 is assumed to have the same emissivity as AMSU channel 15 (both 89 GHz).
- MHS channel 5 is taken as provided by RTTOV.
- MHS channels 2-4 (in range 157-183 GHz) are assumed to have the same emissivity as channel 5 (190 GHz).

This means there are at most 5 independent values which define the emissivity in all 20 AMSU+MHS channels (and hence at most 5 eigenvectors are needed to be fitted in the retrieval).

For IASI, RTTOV uses the Borbas/ University of Wisconsin emissivity database RD-10This is based on a set of 416 eigenvectors of the measured emissivity of a set of natural materials, defined on a 416 point spectral grid spanning a spectral range of 699.3 to 2774.30 cm^{-1}. Emissivity values are extrapolated at fixed value for the channels in the CO2 band below 699.3 cm^{-1}. Spatial distributions of emissivity are determined by fitting these eigenvectors to MODIS observations. Only the first 6 spectral patterns are used for this (due to the limited number of MODIS channels), so only these spectral patterns are represented in the emissivity atlas.

By deriving emissivity patterns from the RTTOV atlases, we can therefore only obtain up to 6 characteristic spectral patterns, together with realistic estimates of their variability to use in the *a priori* covariance. A few more patterns might be expected from the joint IASI+MW covariance (potentially adding 5 degrees of freedom), however in practise there is substantial correlation, and we therefore only take the first 6 spectral patterns from this matrix. IASI retrievals are affected by additional spectral patterns which are not represented in the atlases. In order to address this, further patterns from the set of 416 Wisconsin eigenvectors are added to the set of spectra to be fitted as follows:

1. The 416 Wisconsin patterns, in 416x416 matrix **M,** are interpolated onto the spectral sampling of IASI used in the retrievals (139 channels), defined in 139x416 matrix $\mathbf{M_I}$.
2. The first six Eigenvalues corresponding to the $\mathbf{M_I}$ can be obtained from the variability of the associated patterns in the RTTOV climatology. The Eigenvalues associated with the Wisconsin patterns are not known however it is assumed that they should decrease in magnitude in the order in which they are provided by Wisconsin. For what follows it is mainly important that the order of the additional patterns is maintained (so the most likely spectral variations remain occur first in the final set of patterns). We assume that each pattern from number 7 onwards has an eigenvalue which is 1.3 times smaller than the previous pattern.

3. Having defined the Eigenvectors $\mathbf{M_I}$ and associated Eigenvalues $\boldsymbol{\omega_I}$. A new set of patterns which are orthogonal to the original six are obtained as follows
4. Each pattern (column i=1,416 of $\mathbf{M_I}$) is scaled by the square root of its Eigenvalue to obtain $\mathbf{p_i}$
5. The six original patterns are fitted to $\mathbf{p_i}$ to obtaining the residual pattern

$$\mathbf{p_i}' = \mathbf{p_i} - f(\mathbf{p_i}, \mathbf{R_I})$$

Equation 9

where $f(\mathbf{p_i}, \mathbf{R_I})$ is a simple least squares fit of the six RTTOV based patterns (for the IASI channels), $\mathbf{R_I}$, to $\mathbf{p_i}$.

6. A new spectral covariance is constructed from all 416 residual patterns ($\mathbf{p_i}$' for i=1,N).
7. A new set of Eigenvectors, with appropriately ordered Eigenvalues are obtained by decomposing this covariance matrix to obtain $\mathbf{M_I'}$.
8. The patterns in $\mathbf{M_I'}$ are added to the 6 RTTOV based patterns, to obtain the full set **M'** (for both MW+IASI). Elements corresponding to MW channels are assumed to be 0. The combined list of Eigenvalues is also obtained, **ω'**.
   In principle this results in a list of 422 patterns, however many of these have numerically negligible Eigenvalue. Only a limited number of these patterns are fitted in the retrieval. The retrieval fits the weights (defined in vector **v**) of each pattern such that the emissivity modelled in RTTOV, **e**, is given by

$$\mathbf{e} = \mathbf{v}\,\mathbf{M'}$$

Equation 10

The *a priori* errors for each element of **v** are assumed to be the minimum of the square root of the corresponding Eigenvalue in ω' or 0.01 (values smaller than this are not allowed to prevent too tight a prior constraint). The *a priori* covariance is assumed to be diagonal. A *priori* values for **v** are set by fitting the chosen set of patterns to the standard emissivity given by the RTTOV emissivity atlas for a given scene. Differences between this fit and the RTTOV predicted emissivity are small.

A number of options have been investigated as follows:

- Correlations between IASI and MW emissivities can be included or not. In the latter case the first six patterns are determined independently from the IASI channels. Then the first 5 patterns are taken from the RTTOV covariance for the MW channels. Then additional patterns for IASI are added to the list. The user controls the number of IASI patterns fitted the 5 MW patterns are counted in addition, so in results below a retrieval which indicates 20 patterns always includes 20 patterns relevant to the IASI range. If MW correlations are not included then an additional 5 patterns will be included for the

MW (25 patterns in all). Note that switching on or off the correlation affects the first 6 patterns in the IASI range, which in turn changes all patterns affecting the IASI range.

- In test retrievals over desert (which are particularly prone to high cost due to presumed issues with RTTOV emissivity), it was noted that fit residuals could be significantly reduced if a pattern representing the first spectral derivative of the Wisconsin mean IASI emissivity spectrum was included in the fit (this has quite sharp gradients in the 10 micron region). This effectively corresponds to fitting a wavelength shift of the mean emissivity. Residuals improve further if pattern representing a wavelength stretch of the mean emissivity is included. We therefore include both these patterns in the main simulations reported here. These are inserted into **M'** in order after the RTTOV based patterns, before the additional IASI patterns discussed above.
- Tests have been run with 10, 20 or 30 patterns (columns of **M'**) fitted (including 6 RTTOV patterns and the 2 wavelength shift related patterns).

In the version 1 scheme, MW/IR correlations are assumed and 20 patterns are used (including the spectral shift pattern). Emissivity from the RTTOV atlas is used to defined the prior state itself, via fitting the retrieval Eigenvectors to the RTTOV estimate of emissivity for the given scene (analogous to the approach used to defined the first guess state for profiles described in the previous section).

The spectral patterns used are illustrated in Figure 2-7.

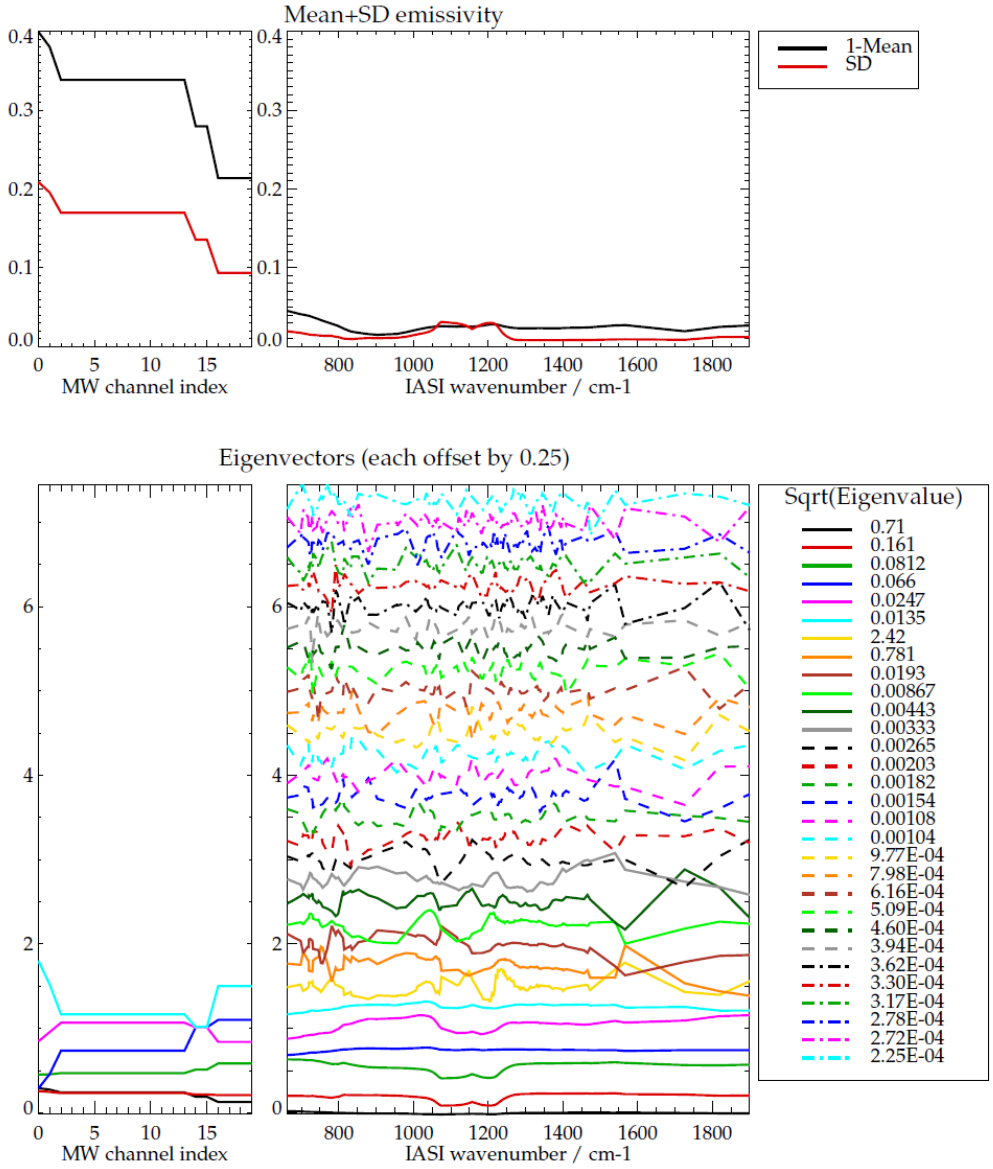

Figure 2-7: Bottom panels shows first 30 spectral patterns used to represent surface emissivity in the retrieval. Each eigenvector is shown offset by 0.25 with respect to the previous vector (for clarity). Only non-zero MW Eigenvectors are shown. The top panels show the mean and standard deviation of the emissivity (note 1 minus the mean emissivity is shown).

2.7.4 Cloud

The area fraction and top pressure of a black-body cloud are both retrieved.

For cloud fraction, the state vector is the natural logarithm of cloud fraction with prior and first guess value $\ln(0.01)$ and prior error 10. The log representation is adopted to prevent negative values of cloud fraction arising.

For cloud top height, the state vector is “z-star” of the cloud pressure. “z-star” is pressure converted to approximate altitude in km:

$$z^* = 16\,(\,3 - \log_{10} p\,)$$

Equation 11

The prior and first guess value is assumed to be 5km with *a priori* error also 5km.

2.7.5 Bias correction patter weights

Weights for the systematic bias correction patterns (section 2.5.1.3) are retrieved. The prior value for each weight is assumed to be zero, with a prior error of 1.

2.8 Characterisation of uncertainty and vertical sensitivity

The error covariance of the solution from an OEM retrieval is given by

$$\boldsymbol{S_x} = \left(\boldsymbol{S_a^{-1}} + \boldsymbol{K^T S_y^{-1} K}\right)^{-1}$$

Equation 12

Where $\boldsymbol{K}$ is the weighting function matrix which contains the derivatives of the FM with respect to each element of the (solution) state vector. The square-roots of the diagonal elements of this matrix are referred to as the *estimated standard deviation* (ESD) of each element of the state vector.

For IMS, the transformation from the state vector to derived profiles, is expressed as a matrix operation (Equation 7). Given the error covariance on the state (from the OEM equations), corresponding error covariance for the layer averages is obtained (e.g. for temperature) by

$$\boldsymbol{S_T = M_T\, S_{x:T} M_T^{\,T}}$$

Equation 13

Where $\boldsymbol{S_{x:T}}$ here is the (square) sub-matrix of the full $\boldsymbol{S_x}$ for the temperature state vector elements only.

For water vapour and ozone, the additional coversion from log units is required to obtain the covariance of the mixing ratio profile in ppmv:

$$\boldsymbol{S_W = (ww^T)\, M_W\, S_{x:w} M_W^{\,T}}$$

Equation 14

Where $\boldsymbol{w}$ is the retrieved water vapour profile in ppmv (from Equation 8).

$\boldsymbol{S_x}$ can be considered in two terms:

$$\boldsymbol{S_x = S_n + S_s}$$

Equation 15

Where $\boldsymbol{S_n}$ describes the retrieval noise, the uncertainity directly related to the specified measurement errors (characterised by $\boldsymbol{S_y}$)**.** $\boldsymbol{S_s}$ describes the smoothing error, i.e. departure from the true state related to variability in the state (expressed in $\boldsymbol{S_a}$) to which the measurements are not sensitive.

$\boldsymbol{S_n}$ is evaluated using retrieval gain matrix (which gives the sensivity of the retrieval to perturbations in the measurement):

$$\boldsymbol{G = S_x K^T S_y^{-1}}$$

Equation 16

$$\boldsymbol{S_n = G^t S_y G}$$

Equation 17

The corresponding error covariance for the layer averages is given by

$$\boldsymbol{S_{Tn} = M_T\, S_{n:T} M_T^{\,T}}$$

Equation 18

The standard deviations of the noise on a derived layer-averages, ΔT_n, are given by the square-root diagonals of this matrix.

Issues relating to the vertical sensitivity of the retrieval and influence of the prior can be addressed using the averaging kernel which characterises the sensitivity of the retrieved state to the actual state (e.g. for temperature):

$$\boldsymbol{A_{f:T} = GK_{f:T}}$$

Equation 19

Where weighting function $\boldsymbol{K_f}$ is distinguised from $\boldsymbol{K}$ in that derivatives are computed with respect to perturbations on fine atmospheric grid, $\boldsymbol{p_{atm}}$ (as opposed to the state vector). In IMS this fine grid is defined by the 101 RTTOV pressure levels. This matrix is not square and has dimensions number of Eigenvector weights in the state vector by 101 „true“ profile pressure levels.

The (square) averaging kernel derived using the state vector weighting function ($\boldsymbol{A = GK}$) would give the derivative of the retrieved Eigenvector weights with respect to true profile perturbations with shape given by the Eigenvectors. These are of relatively little practical use and averaging kernels in this form are only used in IMS to obtain the degrees of freedom for signal (DOFS) for specic retrieval products. This is given by the trace (sum of the diagonal elements) of the sub-matrix of $\boldsymbol{A}$ corresponding to a specifc product.

The averaging kernel of retrieved temperautre profiles (defined on the RTTOV levels) with respect to perturbations on the fine atmospheric grid a layer-average is

$$\boldsymbol{A_{Tf} = M_T A_{xf:T}}.$$

Equation 20

For water vapour, the averaging kernel for the retrieved profile in ln (ppmv), with respect to perturbations in the true profile is given by

$$\boldsymbol{A_{Wf} = M_{Wf} A_{xf:W}}.$$

Equation 21

Where $\boldsymbol{A_{xf:W}}$ is the averaging kernel for the water vapour state vector elements with respect to perturbations in ln(ppmv) on the fine atmospheric levels.

Note that the RTTOV 101 level pressure grid spans the range 0.005 to 1100 hPa, i.e. from below surface to approximately 84km altitude. RTTOV will cut-off the defined profiles at the

defined input surface pressure. Averaging kernels will be by definition zero for true perturbations below the profile-specific surface pressure.

2.9 Comparing retrievals to independent profiles

The consistency of an independent profile with a retrieved profilecan be tested using the averaging kernels. E.g for temperature:

$$\boldsymbol{T_{ind:AK}} = \boldsymbol{m_T}(\lambda) + \boldsymbol{A_{Tf}} \left(\boldsymbol{T_{\mathbf{ind}}} - \boldsymbol{m_T}(\lambda) \right)$$

Equation 22

A similar equation holds for water vapour and ozone, though it should be noted that $\boldsymbol{m_T}(\lambda)$ is defined in ln(ppmv) and $\boldsymbol{A_{Tf}}$ similarly derivatives of retrieved log(ppmv) profiles with respect to perturbations in the true ln(ppmv). $\boldsymbol{W_{\mathbf{ind}}}$ should therefore also be given in units of ln (ppmv). The exponent of the resulting $\boldsymbol{W_{ind:AK}}$ will give the estimated profile in ppmv.

Considering a sufficiently large set of independent profiles, differences between $\hat{T}$ and $T_{ind:AK}$ is expected to be consistent with the estimated retrieval noise matrix, $\boldsymbol{S_{Tn}}$. Differences between the retrieved profile and that of the independent data itself should be consistent with the total solution covariance, $\boldsymbol{S_{Tx}}$ (provided $\boldsymbol{S_a}$ is consistent with the true variability and this is also consistently represented in the set of independent profiles).

3. L2 PRODUCT

3.1 Overview

L2 products are output in CF-compliant NetCDF format. IMS can potentially generate a huge amount of output data, particularly if all retrieval diagnostics (averaging kernels, covariances matrices etc) are to be included. The following choices are made as a trade-off between simplifying the use of the products while maximising the amount of information retained without generating excessively large products:

- IASI orbits are divided into granules of 50 scan lines for processing. L2 files are reported on this basis.
- Only scenes which pass the brightness temperature difference test are processed (see section 2.5.3), however some basic information is reported for all scenes (including the calculated brightness temperature difference and geolocation)
  - Profiles are reported, with their estimated errors (total and noise), on the 101 RTTOV pressure levels (despite being retrieved in terms of fewer principal components). These are stored as short integers with defined scale/offset values to convert to physical units. Where reported values are equal to defined valid_min/valid_max this should be taken to indicate that the true value (after applying the given scale factor and offset) has exceeded the range allowed by the data type used to store the values. Scale factors and offset values used are chosen such that this will rarely happen.
- Emissivity is reported as the spectrum sampled at the same spectral points used in the retrieval.
- The full state vector is also including in full floating point representation. If necessary profile values can be reconstructed using the state vector values in combination with the stored Eigenvectors.
- Full retrieval diagnositcs (averaging kernels, and covariances) are only reported for a sub-set of scenes. The file format is defined such that the sampling can be differented for temperature, water vapour and ozone, however in the version 1 processing the same sampling is chosen for all three profile products. In this case, diagnostics are reported for 1 pixel in 4 along and across-track (i.e. 1 in 16 pixels overall). This gives approximately uniform spatial sampling along/across-track (before considering the cloud mask). A quasi-regular sampling of the across-track scan is adopted so that each of the 4 IASI detector pixels are approximately evenly sampled (a strictly regular 1 in 4 sampling would always take the same detector pixel). In addition,

diagnostics are reported for all scenes within 200km of WOUDC ozone sonde stations.

- When reported, averaging kernels are in the form $\boldsymbol{A_{xf}}$ (see section 2.9), i.e. for state vector elements (Eigenvector weights) with respect to perturbation in the true profile on the 101 level RTTOV grid. To reduce their size, they are only give for every other level in the 101 level grid, i.e. at 51 levels. Since averaging kernels are intrincically related to the grid on which true profiles are defined, a user must interpolate the kernels (in the „true“ dimension) from the 51 sub-set to the full 101 level grid before using them (or otherwise make allowance for the thickness of the layers associated with the kernels).
- When reported, the solution covariance and noise covariance are given only for the profiles separately (i.e. correlation between profiles are not retrained) and in a form which avoids storing the identical elements on both sides of the diagonal (these covarianes are by definition symetric). Each N x N covariance is stored as a vector of N*(N+1)/2 elements, comprising first the diagonal elements, then the 1st superdiagonal elements, then the 2nd superdiagonal elements etc.
- The „prior“ profiles for temperature, humidity and ozone are given tabulated as a function of latitude, not for every scene.
- The (fixed) Eigenvectors which relate the state to the full profile are also given so that the user can reconstruct averaging kernels and covariances for the profiles on the full pressure grid.
- ECMWF analysis profiles (from ERA-interim) are also reported for a sub-set of scenes. These are used to define the first guess state so are readily available when the processor is run.Outputting them to the L2 files enables straightforward monitoring of the L2 product performance.
- The measurement vector, brightness temperature difference spectra (at first guess) and fit residuals (at solution) are also given for only a sub-set of scenes. In version 1 output this again corresponds to the same sub-set of scenes for which profile averaging kernels are reported. Not that in the retrieval the measurement vector is defined in radiance units. In the output files values are converted to brightness temperature (i.e. the tempature of a balck body which would emit the same spectral radiance at a given wavelength). This is done partly mainly to enable more efficient compression: Spectra are stored as short integers with wavelength independent scale and offset.

3.1.1 L2 File name

An example L2 file name is L2 file name is

```
ral-l2p-tqoe-iasi_mhs_amsu_metopa-tir_mw-
20070701203256z_20070701221456z_750_799-v0210.nc
```

Which contains the following tags, separated by "-":

- [origin]: Indicator of where the data was processed, in this case "ral" for RAL.
- [processing level]: In this case "l2p"
- [retrieved species]: In this case "tqoe" for temperature, humidity, ozone and emissivity.
- [instrument]: In this case "iasi_amsu_mhs"
- [platform]: In this case "metopa" for Metop A
- [algorithm id]: In this case "tir_mw", indicating comined use of TIR and MW measurements
- [start date/time] in format YYYYMMDDhhmmss (UT). This is identical to date/time used in the original L1 file (for a full orbit).
- [end date/time] (as above)
- [scanline sub-set start]: This contains the start scanline index index within the L1 orbit of L2 processed data.
- [scanline sub-sent end]: (as above)
- [version ID]: Version ID for the product (v0210).

3.2 Format and content

3.2.1 Global attributes

- title: RAL IASI/AMSU/MHS (IMS) Temperature/Humidity/Ozone/Emissivity/Cloud Product
- project: NCEO
- licence: The data providers [STFC-RAL Remoe-Sensing Group] request that: 1. RAL Remote-Sensing Group be consulted on appropriateness of intended data usage prior to publication of results in any refereed journal or international conference paper and 2. Co-authorship or acknowledgement to RAL Remote-Sensing Group be invited in any refereed journal publication or international conference paper in which the data, or results derived from them, are reported.

- sensor: IASI/MHS/AMSU
- funding: RAL IMS products are developed with funding from the UK National Centre for Earth Observation
- institution: STFC Rutherford Appleton Laboratory
- filename: The L2 filename
- input_filename: The L1 file name processed
- processing_flags: List of processing flags set in the processing job.
- conventions: CF-1.6

3.2.2 Dimensions

- npi: Number of input scenes (i.e. individual IASI measurements in the processed L1 granule)
- npres: Number of scenes for which retrieval resoluts are reported. In the version 1 data, this is all scenes for which pas the brightness temperature difference test (see section 2.5.3)
- npiak_X: Number of scenes for which averaging kernels are reported for profile product X, where X can be “w” for water vapour, “t” for temperature or “o” for ozone
- npisx_X: As above for solution covariance matrices.
- nnwp: Number of scenes for which NWP (i.e. ECMWF ERA-interim) profiles are reported.
- nbtds: Number of scenes for which measurements and brightness temperature difference spectra are reported (i.e. difference between measurement and simulation based on first guess state)
- nresid: Number of scenes for which solution fit residual spectral are reported.
- nbcf=2: Number of bias correction patterns fitted.
- nz=101: Number of RTTOV pressure levels (and pressure levels of the reported profiles)
- nXpc=28: Number of Eigenvectors used to represent profile product X, where X can be “w” for water vapour, “t” for temperature or “o” for ozone or “em” for emissivity.
- nlb: 18
- nsp=159: Number of spectral points (AMSU+MHS+IASI)
- naks=51: Number of “true” profile levels for which averaging kernels are reported, corresponding to every other level of the RTTOV grid.
- navhrr=6: Number of Metop AVHRR channels
- nx=82: Total number of elements in the state vector.
- nvsxt=406: Number of stored covariance values for temperature (per profile)
- nvsxw=171: As above for water vapour

- nvsxo: 55: As above for ozone.

3.2.3 Variables

The following bullets give the variable type (upper case), variable name and dimensions (in brackets) of each variable in the L2 files. Sub-bullets define relevant attributes for each variable. Symbol “X” is used to indicate variables which are reported for each profile variable. X can be “w” for water vapour, “t” for temperature or “o” for ozone

- NC_UBYTE do_retrieval (npi):
  - long_name: Flag indicating scenes which pass quality control (BTD flag zero)
- NC_UBYTE do_ak_X (npi)
  - long_name: Flag indicating scenes for profile product “X” for which full diagnostics are provided including averaging kernels
- NC_UBYTE do_sx_X (npi)
  - long_name: Flag indicating scenes for profile product “X” for which solution error covariance are provided
- NC_UBYTE do_resid (npi)
  - long_name: Flag indicating scenes for which fit residual spectra are provided
- NC_UBYTE do_sample (npi)
  - long_name: Flag indicating scenes forming regular sub-sampling of IASI observations
  - comment: All parameters will be output for these scenes if a retrieval is performed (i.e. when do_sample=1 and do_retrieval=1. Some parameters are provided for the sub-sampled scenes also when a retrieval is not performed
- NC_UBYTE do_station (npi)
  - long_name: Flag indicating scenes close to one of a list of predefined ground-stations (WOUDC ozone-sonde stations in version 1 product)
- NC_UBYTE do_btds (npi)
  - long_name: Flag indicating scenes for which measurements and clear sky brightness temperature difference spectra are provided
- NC_UBYTE do_nwp (npi)
  - long_name: Flag indicating scenes for which NWP profiles are provided
- NC_INT index_nwp (npi)
  - long_name: Index to nearest NWP record for each IASI scene (starts at 0).
- NC_SHORT iat (npi)
  - long_name: Along-track index – scanline within the current granule (from 0-49)

- NC_SHORT ixt (npi)
    - long_name: Across-track index (from 0-119)
- NC_FLOAT jx (npres)
    - long_name: State vector component of cost
- NC_FLOAT jy (npres)
    - long_name: Measurement component of cost
- NC_FLOAT p (nz)
    - long_name: Pressure grid for retrieved profiles (i.e. 101 RTTOV levels)
    - standard_name: air_pressure
    - units: hPa
- NC_SHORT t (nz,npres)
    - long_name: Retrieved atmospheric temperature profile
    - standard_name: air_temperature
    - units: K
    - scale_factor: 0.00625
    - add_offset: 200
- NC_SHORT w (nz,npres)
    - long_name: Retrieved natural logaritm of the atmospheric water vapour profile in parts per million by volume
    - units: ln(ppmv)
    - scale_factor: 0.0003
    - add_offset: 6
- NC_SHORT o (nz,npres)
    - long_name: Retrieved natural logaritm of the atmospheric ozone profile in parts per million by volume
    - units: ln(ppmv)
    - scale_factor: 0.0002
    - add_offset: -0.5
- NC_UBYTE t_err (nz,npres)
    - long_name: Error in retrieved atmospheric temperature profile
    - standard_name: air_temperature
    - units: K
    - scale_factor: 0.05
    - add_offset: 0
- NC_UBYTE w_err (nz,npres)
    - long_name: Error in retrieved natural logarithm of the atmospheric water vapour profile
    - units: 1
    - scale_factor: 0.0025

    - add_offset: 0
- NC_UBYTE o_err (nz,npres)
    - long_name: Error in retrieved natural logarithm of the atmospheric ozone profile
    - units: 1
    - scale_factor: 0.004
    - add_offset: 0
- NC_SHORT X_nwp (nz,nnwp)
    - long_name: NWP profile for product X
    - units: depends on X
    - scale_factor: depends on X
    - add_offset: depnds on X
- NC_FLOAT xX (ntpc,npres)
    - long_name: State vector for profile product X in terms of principal component weights
- NC_FLOAT tsk (npres)
    - long_name: State vector for surface temperature
    - standard_name: surface_temperature
    - units: K
- NC_FLOAT tsk_nwp (nnwp)
    - long_name: NWP surface temperature
    - standard_name: surface_temperature
    - units: K
- NC_FLOAT cfr (npres)
    - long_name: Retrieved effective cloud fraction
- NC_FLOAT cth (npres)
    - long_name: Retrieved cloud top pressure in z-star scae-height.
    - units: km
- NC_FLOAT bcf (nbcf,npres)
    - long_name: Retrieved bias correction factor
- NC_SHORT em (nsp,npres)
    - long_name: Retrieved emissivity
    - units: 1
    - scale_factor: 2.5e-05
    - add_offset: 0.8
- NC_UBYTE em_err (nsp,npres)
    - long_name: Estimated error on retrieved emissivity
    - units: 1

    - scale_factor: 0.001
    - add_offset: 0
- NC_FLOAT cth_err (npres)
    - long_name: Estimated error on retrieved cloud top height
    - units: km
- NC_FLOAT cfr_err (npres)
    - long_name: Estimated error on retrieved cloud fraction
- NC_FLOAT tsk_err (npres)
    - long_name: Estimated error on retrieved surface temperature
    - units: K
- NC_FLOAT bcf_err (nbcf,npres)
    - long_name: Estimated error on retrieved bias correction factors
    - units:
- NC_FLOAT X_ap (nz,nlb)
    - long_name: Climatological prior for profile product X
    - units: depends on X (K or ln (ppmv))
- NC_FLOAT tsk_ap (nlb)
    - long_name: Climatological prior for surface temperature
    - units: K
- NC_SHORT ak_X (nXpc,naks,npiak_X)
    - long_name: Averaging kernel for
    - units: 1
    - scale_factor: 4.5777e-05
    - add_offset: -0.5
- NC_FLOAT avhrr_radiance (navhrr,npi)
    - long_name: Mean of co-located AVHRR radiances
- NC_FLOAT evecs_X (nz,nXpc)
    - long_name: Eigenvectors to reconstruct profile variable X.
- NC_SHORT conv (npres)
    - long_name: Flag indicating retrieval convergence (1=fully converged). Results with other values should be used with more caution
- NC_INT n_step (npres)
    - long_name: The average number of retrieval steps (number of calls to the forward model)
- NC_SHORT pixel_number (npi)
    - long_name: IASI detector pixel (0-3)
- NC_FLOAT longitude (npi)
    - long_name: Longitude
    - standard_name: longitude

    - units: degree_east
- NC_FLOAT latitude (npi)
    - long_name: Latitude
    - standard_name: latitude
    - units: degree_north
- NC_FLOAT satzen (npi)
    - long_name: Sensor view zenith angle
    - standard_name: sensor_zenith_angle
    - units: degree
- NC_FLOAT solzen (npi)
    - standard_name: solar_zenith_angle
    - long_name: Solar zenith angle
    - units: degree
- NC_UINT sensingtime_msec (npi)
    - long_name: Time of day in msec since midnight (UTC)
    - units: s
- NC_UINT sensingtime_day (npi)
    - long_name: Day of measurement since 00:00h 01-01-2000
    - units: day
- NC_FLOAT sp (npi)
    - long_name: Assumed surface pressure in retrieval
    - units: hPa
- NC_FLOAT ecmwf_altitude (npi)
    - long_name: Estimated surface altitude from interpolated ECMWF ERA Interim data
    - units: m
- NC_FLOAT iasi_altitude (npi)
    - long_name: Estimated surface altitude from interpolated GTOPO30 atlas
    - units: m
- NC_SHORT resid (nsp,nresid)
    - long_name: Residual spectra
    - units: K
    - scale_factor: 0.002
    - add_offset: 0
- NC_FLOAT cwn (nsp)
    - long_name: Wavenumbers associated with measurement vector
    - units: cm^{-1}
- NC_SHORT bt (nsp,nbtds)

    - long_name: Measurement vector in terms of brightness temperature
    - units: K
    - scale_factor: 0.00625
    - add_offset: 200
- NC_SHORT btds (nsp,nbtds)
    - long_name: Brightness temperature difference spectra between observed and simulated measurements for first guess state
    - units: K
    - scale_factor: 0.002
    - add_offset: 0
- NC_FLOAT btd (npi)
    - long_name: Brightness temperature difference between observed and simulated measurements for first guess state, in a window channel at 955.25 cm-1
    - units: K
    - missing_value: -3e+33
- NC_UBYTE btd_flag (npi)
    - long_name: Flag used as indicator of cloud derived from reported variable btd (at 950cm^{-1})
- NC_FLOAT sa_X (ntpc)
    - long_name: Globally averaged a priori covariance for profile product X in terms of PC weights
- NC_FLOAT sa_ts (n1)
    - long_name: Global surface temperature a priori variance
    - units: K^2
- NC_FLOAT vsx_X (nvsxX,npisx_X)
    - long_name: Flattened half of profile product X solution covariance matrix in terms of PC weights
- NC_FLOAT vsxn_X (nvsxX,npiak_X)
    - long_name: Flattened half of profile product X noise covariance matrix in terms of PC weights
- NC_FLOAT X_dofs (npres)
    - long_name: Degrees of Freedom for Signal of profile product X.
- NC_FLOAT em_dofs (npres)
    - long_name: Degrees of Freedom for Signal of surface emissivity
- NC_FLOAT cfr_dofs (npres)
    - long_name: Degrees of Freedom for Signal of Cloud Fraction
- NC_FLOAT cth_dofs (npres)
    - long_name: Degrees of Freedom for Signal of Cloud Top Height

- NC_FLOAT tsk_dofs (npres)
    - long_name: Degrees of Freedom for Surface Temperature
- NC_FLOAT bcf_dofs (npres)
    - long_name: Degrees of Freedom for Bias Correction Factors

3.3 Quality control

The main parameter which can be used for quality control is the cost function at solution. It is recommended that retrievals with total cost function (the sum of variables jx and jy) greater than 1000 are not used. It is possible that for some applications cloud fraction and/or cloud-top-height could be useful as quality control, however quantitative recommendations for using these parameters cannot be given at this time.

4. PRODUCT QUALITY

4.1 Sensitivity and vertical resolution

Some basic retrieval diagnostics from the IMS version 1 data are presented for three typical scenes in Figure 4-1 to Figure 4-3, below. These figures show the following for temperature and water vapour:

- Top-right legend: Gives some basic information about the scene including date, geolocation, scan geometry, together with some out variables (using the names as in the L2 file (see section 3.2.3)
- Left:Comparison of the retrieved (Ret, black), a priori (AP, red), ECMWF analysis (NWP, green) and ECMWF analysis after applying averaging kernels via Equation 22 (NWPxAK, blue). Dashed lines about the retrieved profile indicate the range of the estimated standard deviation (ESD).from the solution covariance.
- Right-hand plots: Show averaging kernels for a regularly sampled subset of the 101 output retrieval levels (each line showing the sensitivity of the given output level to perturbations in the true profile, as a function of z* altitude).

These examples illustrate the capability of IMS to resolve water vapour throughout the troposphere and return profiles close to NWP, even where this deviates strongly from the climatological prior. The altitude up to which the water vapour profile is well resolved follows the tropopuse: Sensitivity to stratospheric water vapour is generally low (due mainly to the relatively low stratospheric mixing ratio).

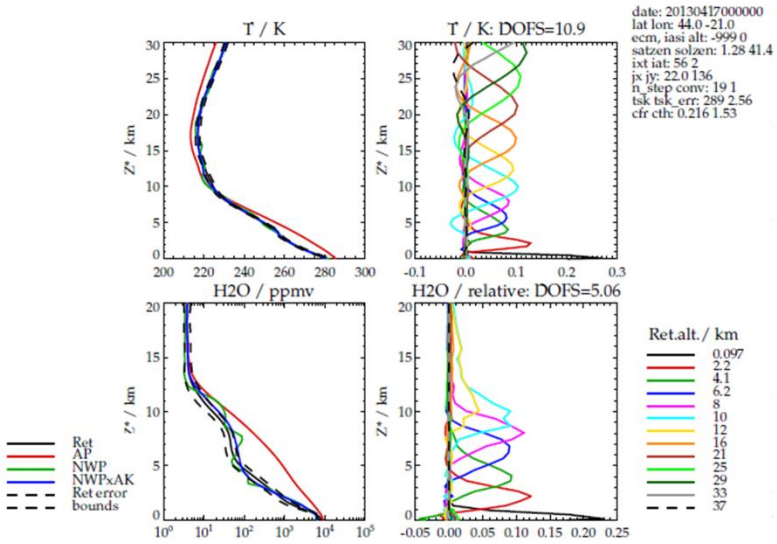

Figure 4-1: Retrieval diagnostics for temperature (top) and water vapour (bottom) for a mid-latitude scene over sea.

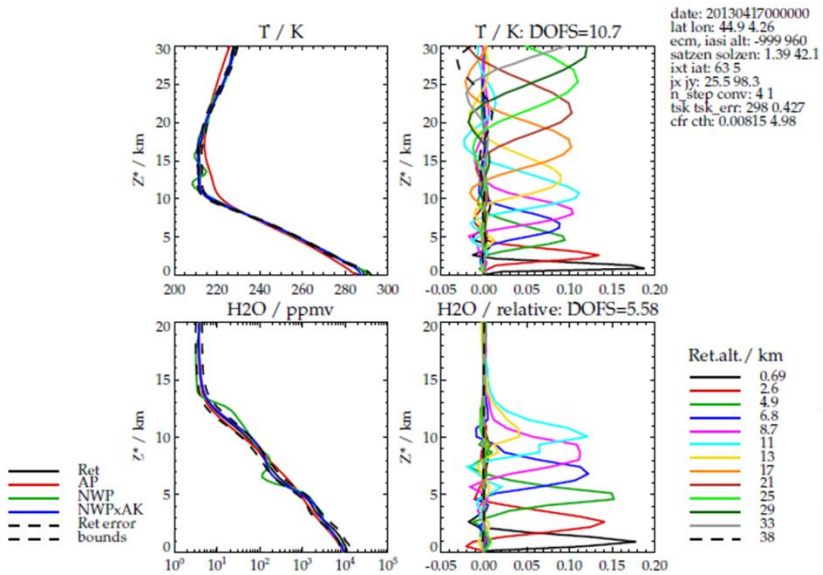

Figure 4-2: Retrieval diagnostics for temperature (top) and water vapour (bottom) for a mid-latitude scene over land.

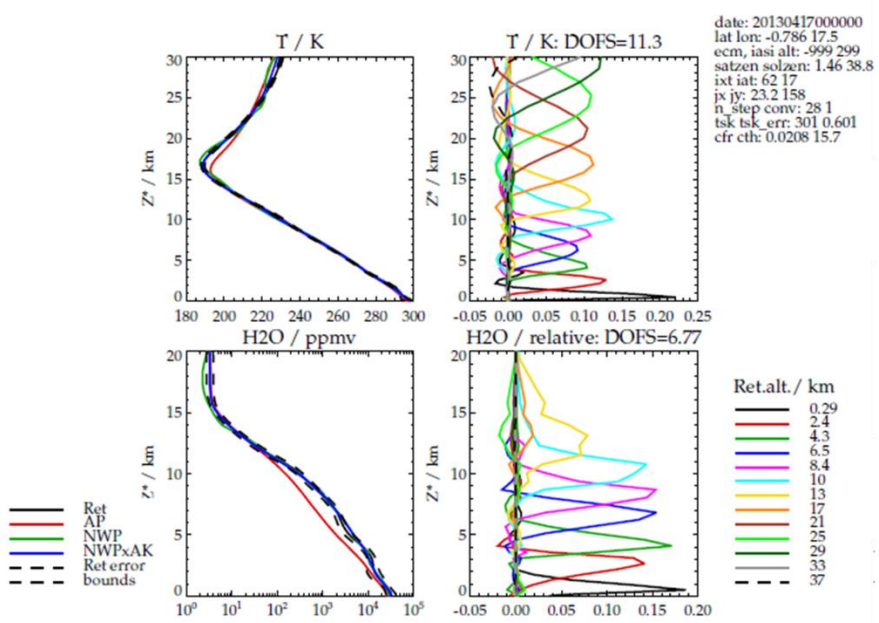

Figure 4-3: Retrieval diagnostics for temperature (top) and water vapour (bottom) for a tropical scene over land.

4.2 Time-series comparisons to ERA-interim

Result from the full time-series of version 1 IMS data are illustrated via Hovmoller diagrams in Figure 4-4 - Figure 4-13. Each figure has six panels showing results at different pressure levels throughout the troposphere up to the UT/LS. Data run from mid 2007 (beginning of the Metop-A operational phase) to the end of 2016. Data in April 2014 is missing due to a break in the availability of data from MHS.

Figure 4-4 shows (mainly for context) the zonal mean time-series of retrieved water vapour. Note that the prior distribution has no time variation, so all the seasonal cycle is derived from the measurements. Throughout results are shown in units of ln(ppmv).

Figure 4-5 shows the estimated standard deviation of solution covariance (averaging into the latitude/time bins). This is an estimate of the total error (noise+smoothing) on the retrieved profiles. Given the log unit, these can be approximately interpreted as relative errors. Within the troposphere errors are at best ~10-20%, with larger uncertainty where the water mixing ratios are smaller (towards higher altitude or polar winter). Errors are particularl large near the surface over Antarctica, due to dry conditions are very low surface temperature.

FFigure 4-6 shows mean differences between IMS and ERA interim. These errors follow the distribution of the ESD to some extent (this would be expected for the smoothing error contribution which may give rise to systematic error).

Figure 4-7 shows differences after taking into account retrieval sensitivity by applying the averaging kernels via Equation 22. Differences are generally much reduced compared to the previou figure, indicating that averaging kernels can explain much of the deviation between ERA-interim and the retrievals. There remains an systematic bias particularly in the upper troposphere, which IMS is 10-20% dryer than ERA-interim.

Figure 4-8 and Figure 4-9 show the standard deviation of the difference between individual retrieved profiles and ERA-interim, before and after accounting for averaging kernels, respectively. These also show the importance of the averaging kernels in explaining much of the differences.

Figure 4-10, Figure 4-11 and Figure 4-12 show the ability of IMS to capture the interannual variability of water vapour. They show the de-seasonalised anomaly i.e. the difference between individual monthly zonal means and the multi-annual zonal mean (considering all years 2007-16) for the same month. Figure 4-10 shows the retrieved time-series while Figure 4-11 and Figure 4-12 show ERA-interim before and after applying averaging kernels. The retrieval tends to smooth the “true” variability towards the higher altitudes, but captures very well both the quasi-random structure and the high anomalies associated with the El Nino in 2015-16 and (to a lesser extent) 2009-10.

Figure 4-13 shows the de-seasonalised anomaly of the difference between IMS and ERA interim with averaging kernels. This could reveal any time dependent instrumental bias (drift, discontinuities) in the retrieval (assuming ERA-interim to be bias free). From this figure it appears that any such errors are low (<<5%).

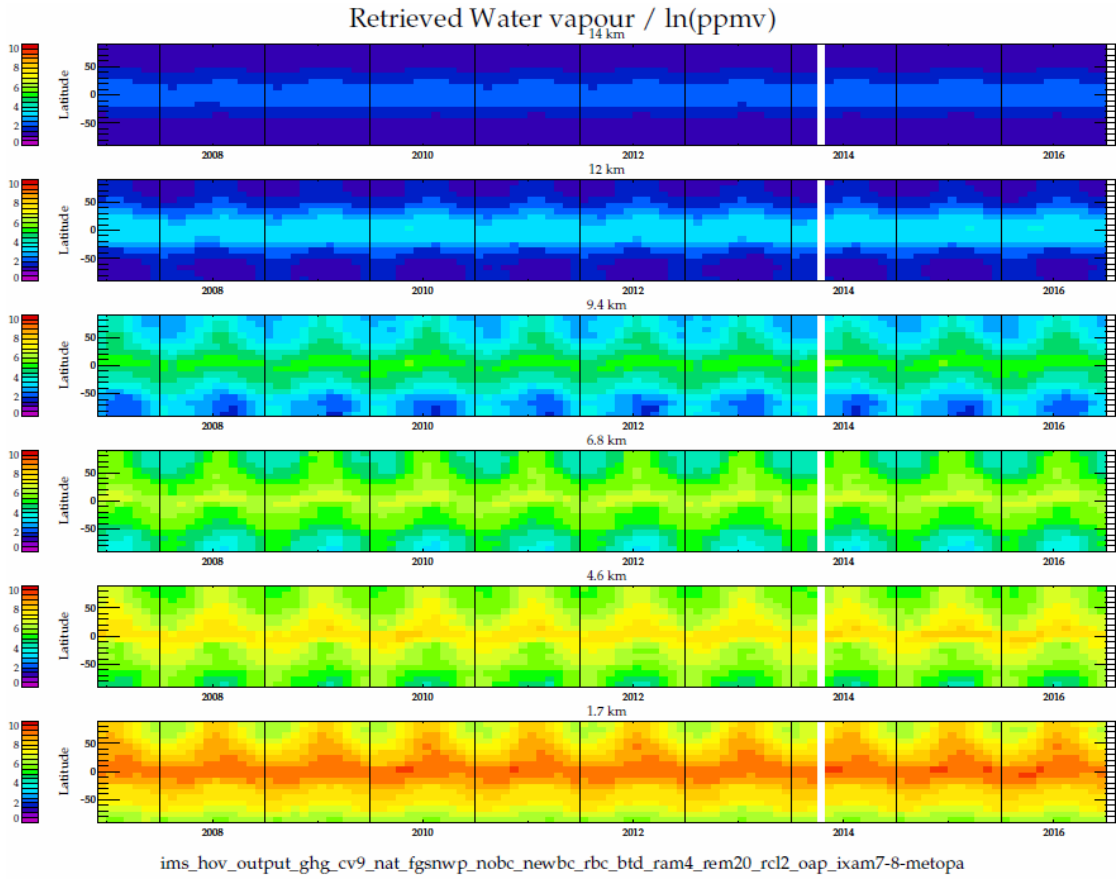

Figure 4-4: Hovmoller plot of retrieved water vapour at six pressure levels from the version 1 processing.

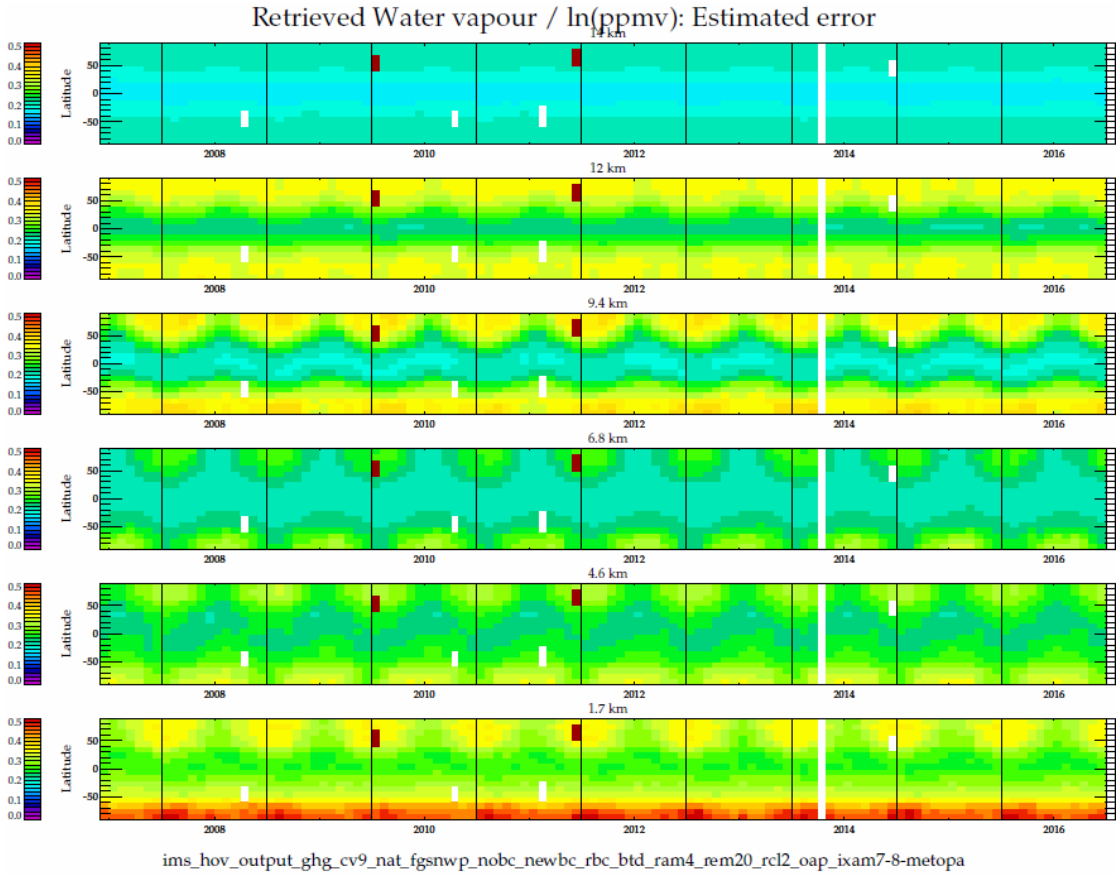

Figure 4-5: Estimated standard deviation of retrieved water vapour at six pressure levels from the version 1 processing.

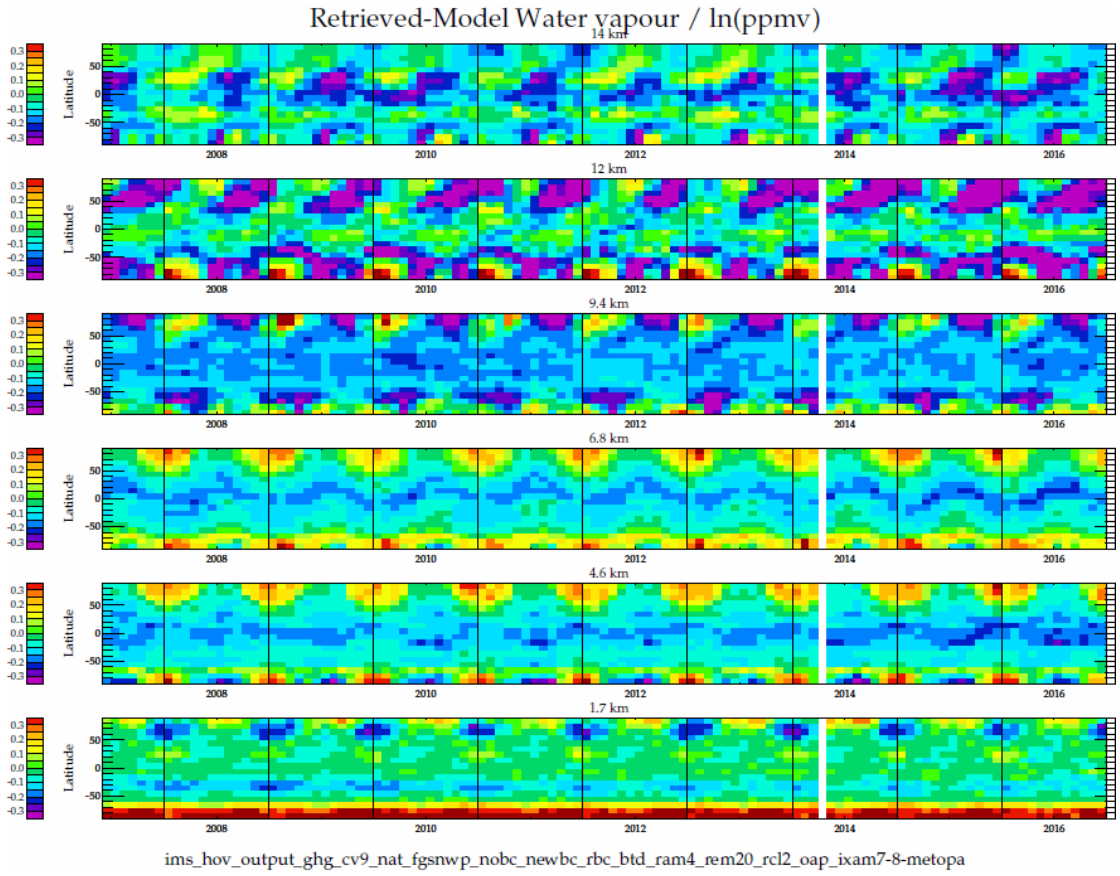

Figure 4-6: Mean differences between retrieved and ERA-interim water vapour, (in ln(ppmv).

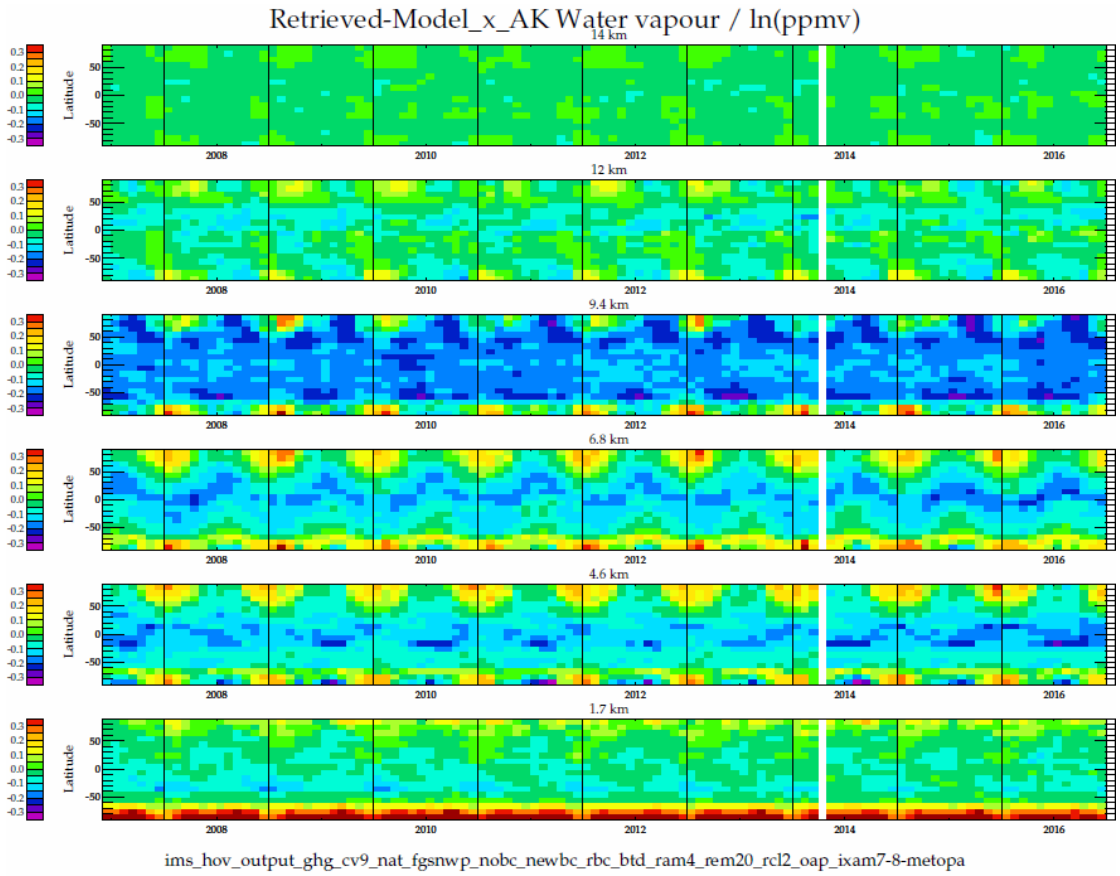

Figure 4-7: Mean difference between retrieved water vapour and ERA-interim, after accounting for averaging kernels.

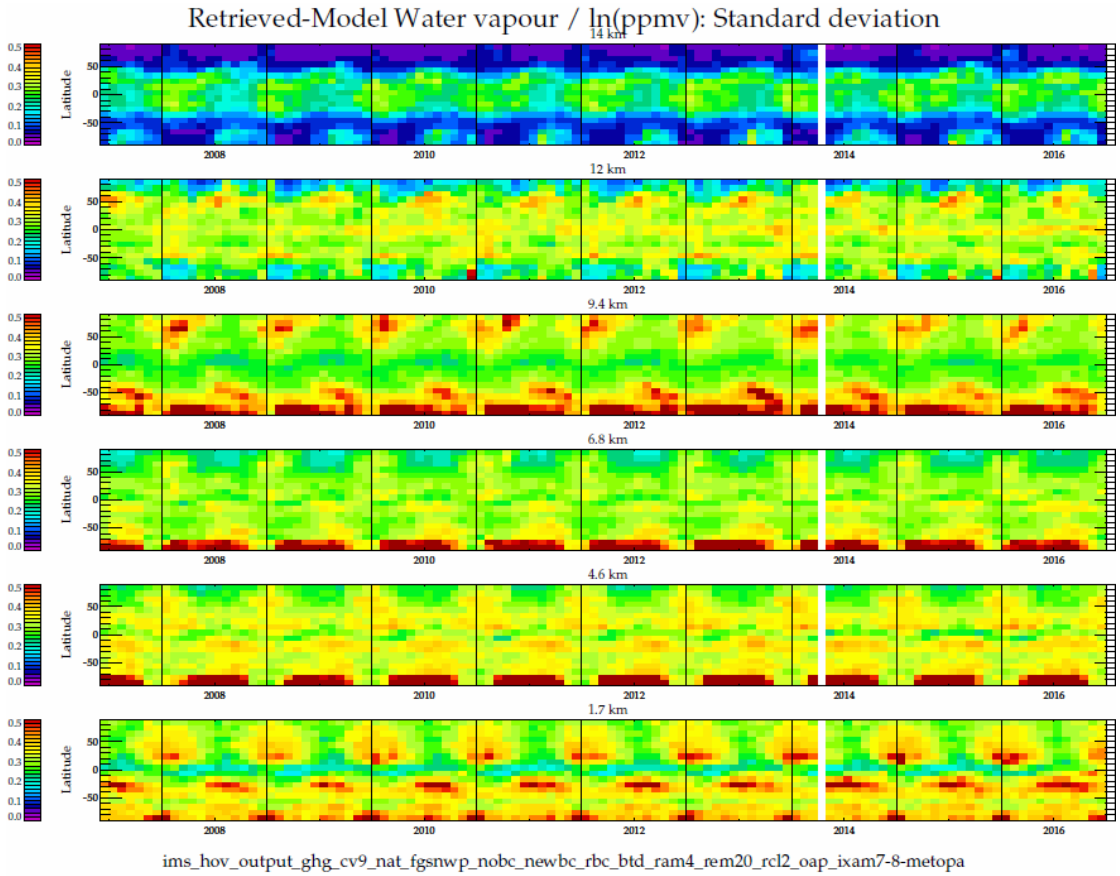

Figure 4-8: Standard deviation of individual profile differences between retrieval and ERA-interim.

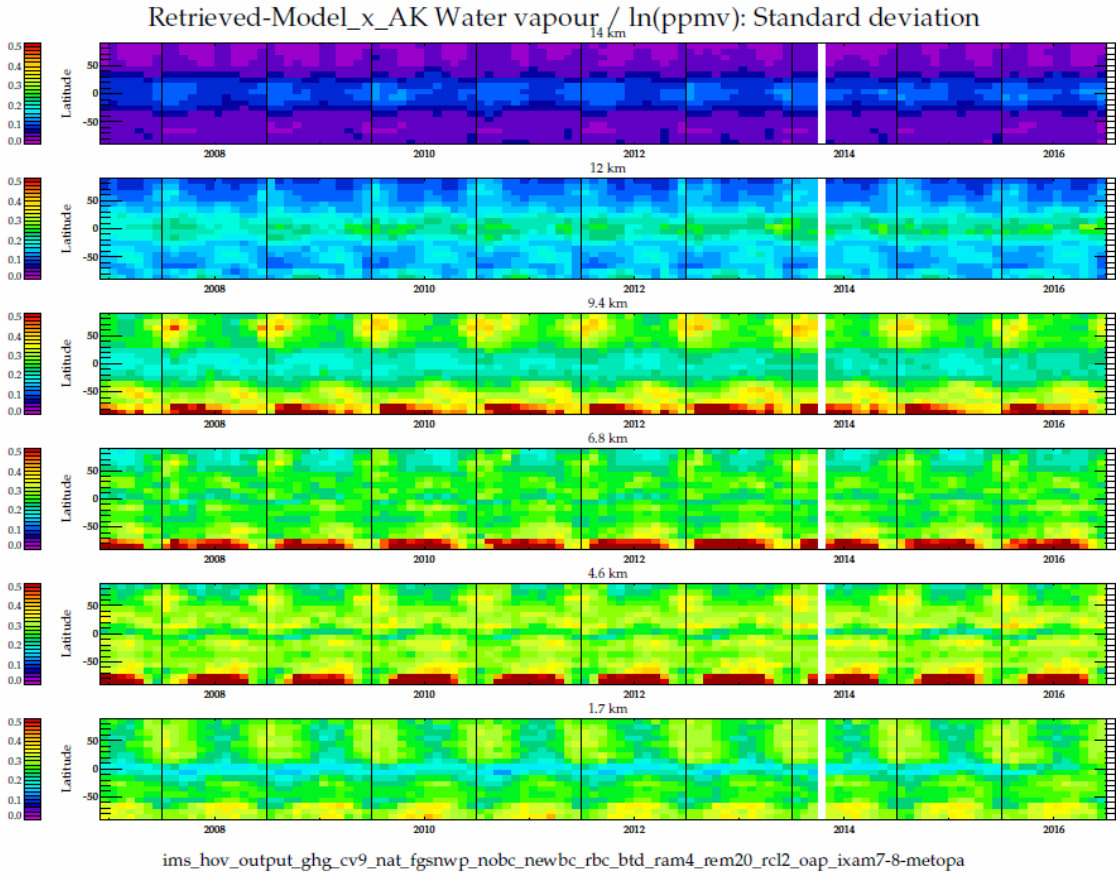

Figure 4-9: Standard deviation of individual profile differences between retrieval and ERA-interim, after accounting for averaging kernels.

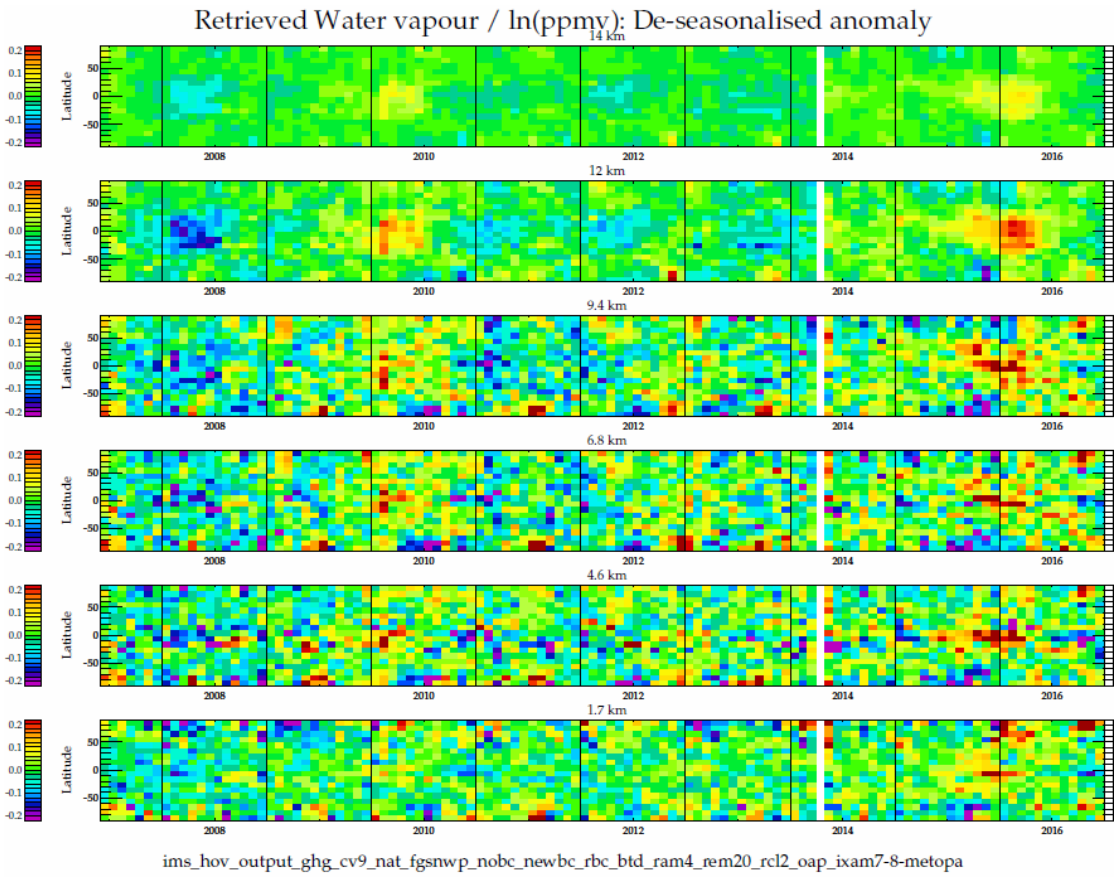

Figure 4-10: De-seasonalised anomaly of retrieved water vapour at size pressure levels from the version 1 processing.

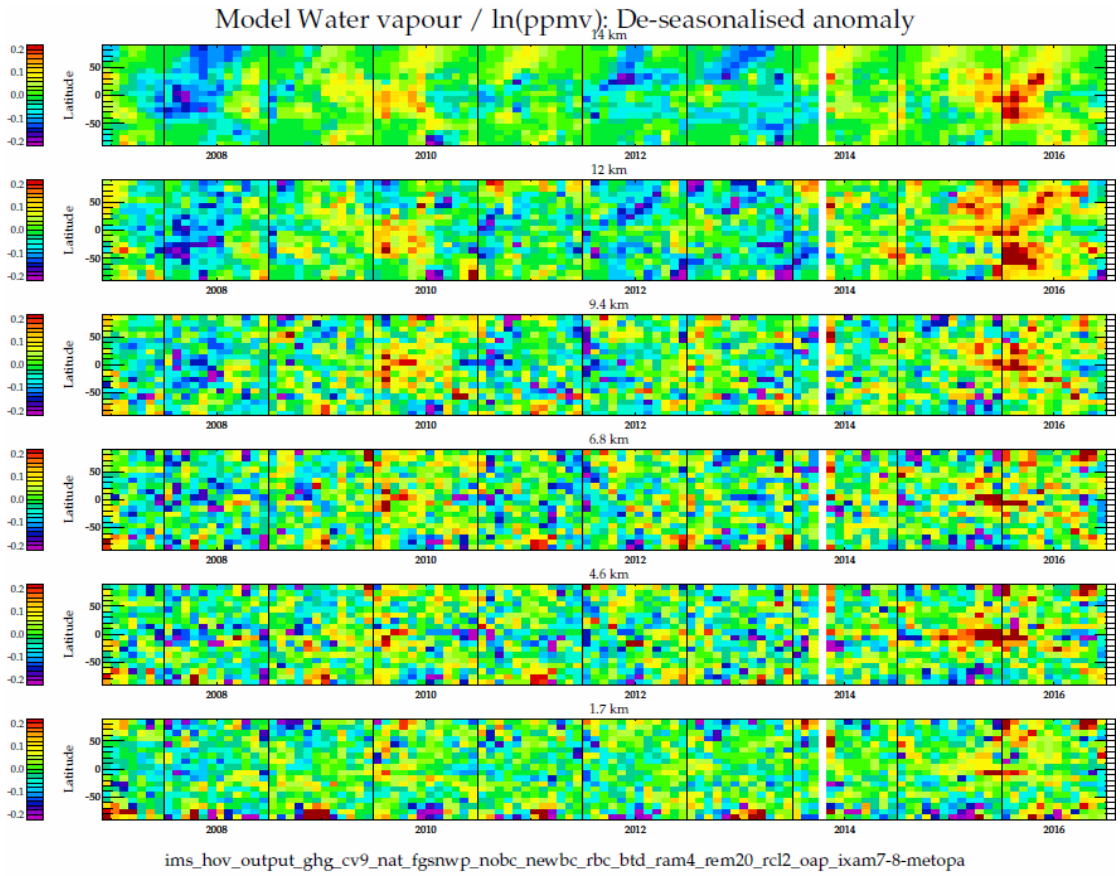

Figure 4-11: De-seasonalised anomaly of ERA-interim water vapour.

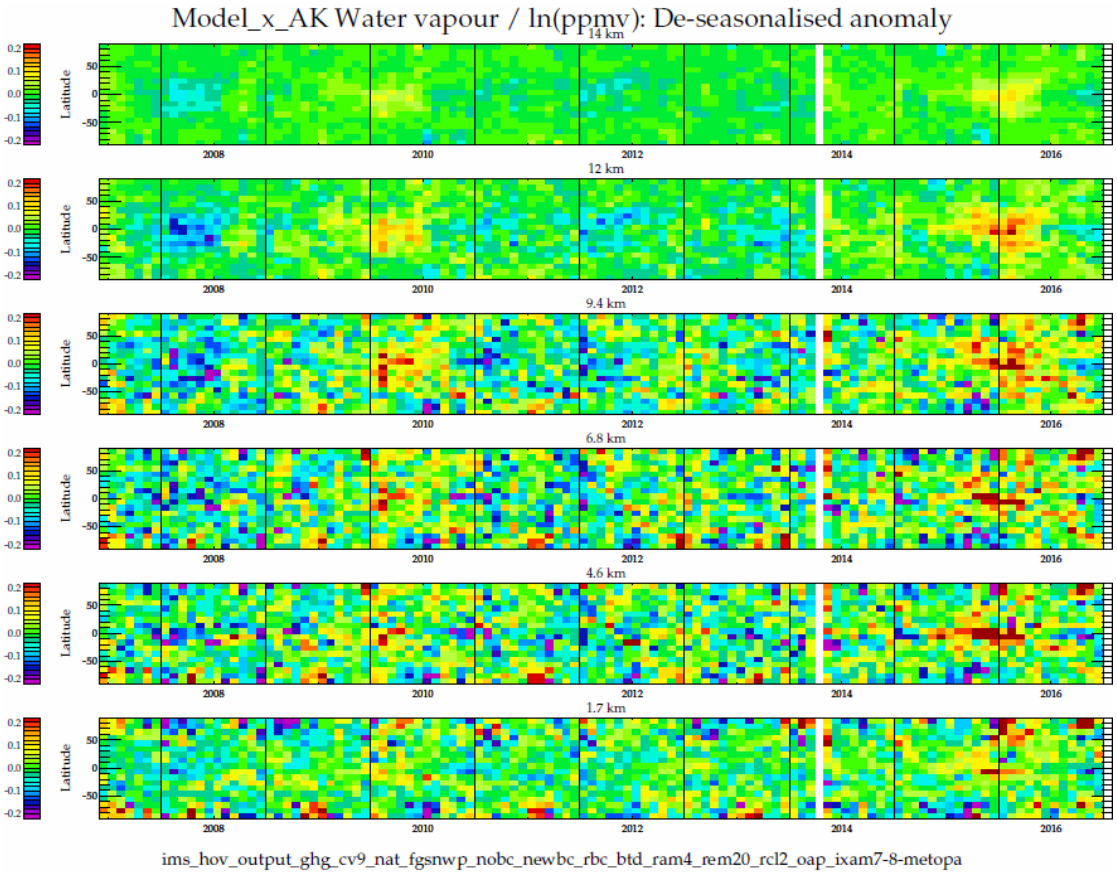

Figure 4-12: De-seasonalised anomaly of ERA-interim water vapour, after accounting for IMS averaging kernels.

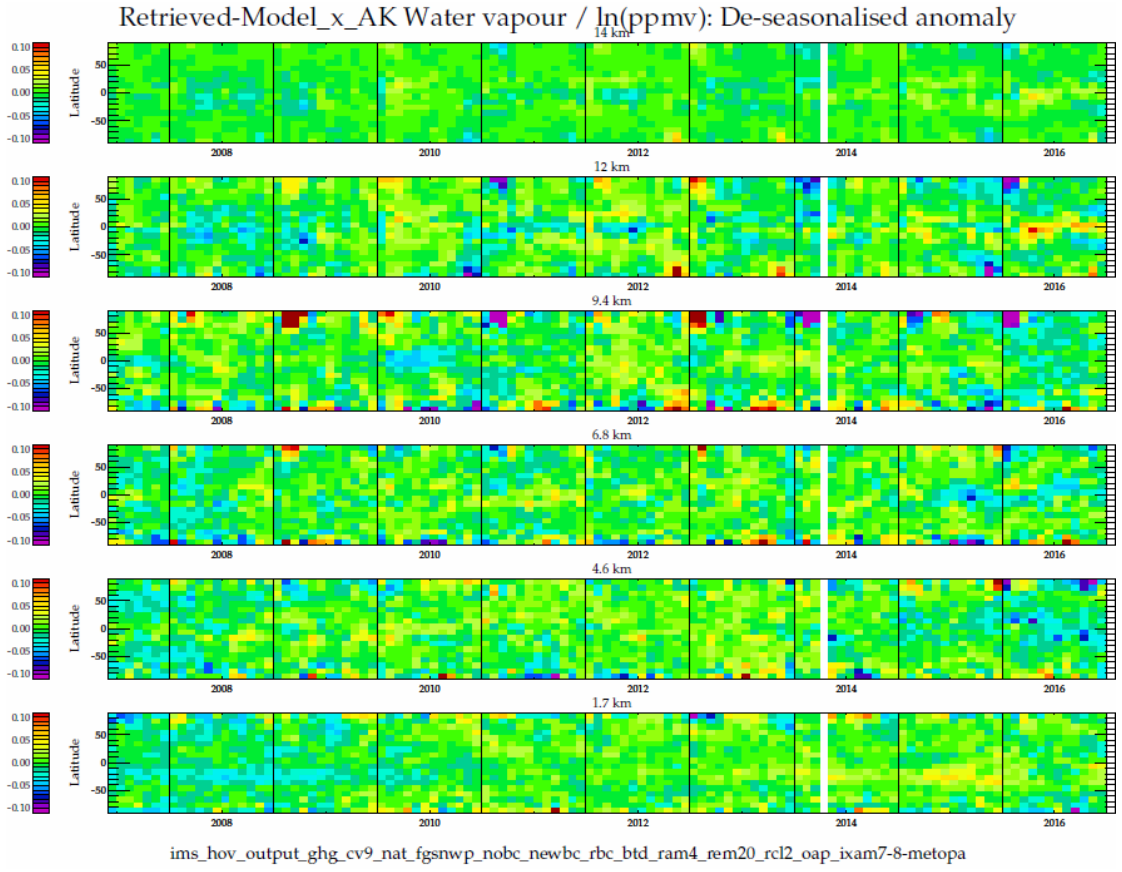

Figure 4-13: De-seasonalised anomaly of the difference between retrieved water vapour and ERA-interim, accounting for averaging kernels..

APPENDIX 1: REFERENCES

RD-1. R. Siddans, D. Gerber, B. Bell, Optimal Estimation Method retrievals with IASI, AMSU and MHS measurements. Final Report, EUM/CO/13/46000001252/THH, 2015

RD-2. Siddans, R., Knappett, D., Waterfall, A., Hurley, J., Latter, B., Kerridge, B., Boesch, H., and Parker, R.: Global height-resolved methane retrievals from the Infrared Atmospheric Sounding Interferometer (IASI) on MetOp, *Atmos. Meas. Tech. Discuss.*, https://doi.org/10.5194/amt-2016-290, accepted for publication Sept'2017.

RD-3. Eumetsat; IASI Level 2: Product Guide; EUM/OPS-EPS/MAN/04/0033; v3E e-signed, 11 July 2017

RD-4. Blumstein, D.; Chalon, G.; Carlier, T.; Buil, C.; Hébert, Ph.; Maciaszek, T.; Ponce, G.; Phulpin, T.; Tournier, B.; Siméoni, D.; Astruc, P.; Clauss, A.; Kayal, G.; Jegou, R. (2004). "IASI instrument: technical overview and measured performances". Proceedings of the SPIE. Infrared Spaceborne Remote Sensing XII 5543: 196–207. doi:10.1117/12.560907.

RD-5. ATOVS Level 1b Product Guide; EUMETSAT; Doc.No. : EUM/OPS-EPS/MAN/04/0030; Issue : v3; Date : 11 January 2010.

RD-6. ATOVS Level 2 Product Guide; Eumetsat; Doc.No: EUM/OPS-EPS/MAN/04/0031; v2, 22 July 2009

RD-7. Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Sci., Hackensack, N.J., 2000.

RD-8. Press, W.H., Teukolksy, S., Vetterling, W.T. and Flannery, B., Numerical Recipes: the art of scientific computing, Second edition, Cambridge University Press, 1995.

RD-9. Siddans, R.; Walker, J.; Latter, B.; Kerridge, B.; Gerber, D.; Knappett, D. (2018): RAL Infrared Microwave Sounder (IMS) temperature, water vapour, ozone and surface spectral emissivity. Centre for Environmental Data Analysis, *date of citation*. doi:10.5285/489e9b2a0abd43a491d5afdd0d97c1a4. http://dx.doi.org/10.5285/489e9b2a0abd43a491d5afdd0d97c1a4

RD-10. Borbas, E. E. and B. C. Ruston, 2010. The RTTOV UWiremis IR land surface emissivity module. NWP SAF report. http://research.metoffice.gov.uk/research/interproj/nwpsaf/vs_reports/nwpsaf-mo-vs-042.pdf

RD-11. Hultberg and August, CANONICAL ANGLES BETWEEN THE IASI OBSERVATION AND FORWARD MODEL SUBSPACES

RD-12. Hultberg and August, THE PIECEWISE LINEAR REGRESSION RETRIEVAL OF TEMPERATURE, HUMIDITY AND OZONE WITHIN THE EUMETSAT IASI L2 PPF VERSION 6

RD-13. Roger Saunders, James Hocking, Peter Rayer, Marco Matricardi, Alan Geer, Niels Bormann, Pascal Brunel, Fatima Karbou and Filipe Aires, RTTOV-10 SCIENCE AND VALIDATION REPORT; NWPSAF-MO-TV-023; Version : 1.11; Date : 23/1/2012

RD-14. Atkinson, N.C, F. I. Hilton, S. M. Illingworth, J. R. Eyre and T. Hultberg; Potential for the use of reconstructed IASI radiances in the detection of atmospheric trace gases; Atmos. Meas. Tech., 3, 991-1003, 2010; https://doi.org/10.5194/amt-3-991-2010

RD-15. Dudhia, A; Zenith Sky Optical Thickness Infrared Spectrum; http://eodg.atm.ox.ac.uk/ATLAS/zenith-absorption

RD-16. Bergamaschi, P., Segers, A., Scheepmaker, R., Frankenberg, C., Hasekamp, O., Dlugokencky, E., Sweeney, C., Ramonet, M., Tarniewicz, J., Kort, E., and Wofsy, S.(2013b): Report on the quality of the inverted CH4 fluxes, MACC-II Deliverable D_43.3, Tech. rep., available at: https://atmosphere.copernicus.eu/documents/maccii/deliverables/ghg/, Joint Research Center, European Commission.

RD-17. D. P. Dee, S. M. Uppala, A. J. Simmons et al, The ERA-Interim reanalysis: configuration and performance of the data assimilation system, QJRMS; 28 April 2011; https://doi.org/10.1002/qj.828

RD-18. Liu, Q, F. Weng, S.J. English, An Improved Fast Microwave Water Emissivity Model, IEEE Transactions on Geoscience and Remote Sensing (Impact Factor: 3.47). 05/2011; DOI: 10.1109/TGRS.2010.2064779

RD-19. Karbou, F., E.Gérard, and F. Rabier, 2006, Microwave land emissivity and skin temperature for AMSU-A and –B assimilation over land, Q. J. R. Meteorol.Soc., vol. 132, No. 620, Part A, pp. 2333-2355(23), doi :10.1256/qj.05.216

APPENDIX 2: GLOSSARY

Term	Definition
AMSU	Advance Microwave Sounding Unit (on Metop)
ATBD	Algorithm Theoretical Basis Document
CCI	ESA Climate Change Initiative
E3UB	End to End ECV Uncertainty Budget
ESA	European Space Agency
FM	Forward Model
IASI	Infrared atmospheric sounding interferometer
IMS	Infra-red Microwave Sounder scheme
IR	Infra-red
MHS	Microwave Humidity Sounder
MW	Microwave
NCEO	UK National Centre for Earth Observation
OEM	Optimal estimation method
PVIR	Product Validation and Intercomparison Report
RAL	Rutherford Appleton Laboratory
RTTOV	Radiative transfer for TOVS (radiative transfer model)
TOVS	TIROS operational vertical sounder (package of MW+IR sounders on operational polar platforms)
UV	Ultra-violet
WOUDC	World Ozone and UV Data centre

End of Document